UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 833

Observational Uncertainties in Water-Resources Modelling in Central America

Methods for Uncertainty Estimation and Model Evaluation

IDA WESTERBERG

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2011

ISSN 1651-6214
ISBN 978-91-554-8090-5
urn:nbn:se:uu:diva-152074

Dissertation presented at Uppsala University to be publicly examined in Alex Hambergsalen, Earth Sciences Centre, Villavägen 16, Uppsala, Friday, June 10, 2011 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Westerberg, I. 2011. Observational Uncertainties in Water-Resources Modelling in Central America. Methods for Uncertainty Estimation and Model Evaluation. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 833. 75 pp. Uppsala. ISBN 978-91-554-8090-5.

Knowledge about spatial and temporal variability of hydrological processes is central for sustainable water-resources management, and such knowledge is created from observational data. Hydrologic models are necessary for prediction for time periods and areas lacking data, but are affected by observational uncertainties. Methods for estimating and accounting for such uncertainties in water-resources modelling are of high importance, especially in regions such as Central America.

Observational uncertainties were addressed in three ways in this thesis; quality control, quantitative estimation and development of model-evaluation techniques that addressed unquantifiable uncertainties. A first step in any modelling study should be the quality control and concurrent analysis of the representativeness of the observational data. In the characterisation of the precipitation regime in the Choluteca River basin in Honduras, four different quality problems were identified and 22% of the daily data had to be rejected. The monitoring network was found to be insufficient for a comprehensive characterisation of the high spatiotemporal variability of the precipitation regime.

Quantitative estimations of data uncertainties can be made when sufficient information is available. Discharge-data uncertainties were estimated with a fuzzy regression for time-variable rating curves and from official rating curves for 35 stations in Honduras. The uncertainties were largest for low flows, as a result of measurement uncertainties and natural variability.

A method for calibration with flow-duration curves was developed which enabled calibration to the whole flow range, accounting for discharge uncertainty and calibration with nonoverlapping time periods for model input and evaluation data. The method compared favourably to traditional calibration in a test using two models applied in basins with different runoff-generation processes. A *post-hoc* analysis made it possible to identify potential model-structure errors and periods of disinformative data. Flow-duration curves were regionalised and used for calibration of a Central-American water-balance model. The initial model uncertainty for the ungauged basins was reduced by 70%. Non-representative precipitation data were found to be the main obstacle to comprehensive regional water-resources modelling in Central America.

These methods bridged several problems related to observational uncertainties in water-balance modelling. Estimates of prediction uncertainty are an important basis for all types of decisions related to water-resources management.

Keywords: Central America, Discharge, Flow-duration curve, Fuzzy regression, GLUE, Model evaluation, Non-stationarity, Observational uncertainty, Precipitation, Quality control, Rating curve, Regionalisation, Uncertainty estimation, Ungauged basins, Water resources.

Ida Westerberg, Department of Earth Sciences, LUVAL, Villav. 16, Uppsala University, SE-75236 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-8090-5

Akademisk avhandling som för avläggande av teknologie doktorsexamen i hydrologi vid Uppsala universitet kommer att offentligen försvaras i Alex Hambergsalen, Geocentrum, Villavägen 16, Uppsala, fredagen den 10 juni, 2011 klockan 10:00. Professor Alberto Montanari från Università di Bologna är fakultetsopponent. Disputationen sker på engelska.

Referat

Westerberg, I. K. 2011. Observationsosäkerheter i vattenresursmodellering i Centralamerika – Metoder för osäkerhetsuppskattning och modellutvärdering. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 833. 75 pp. Uppsala. ISBN 978-91-554-8090-5.

Kännedom om hur hydrologiska processer varierar i tid och rum är grundläggande för hållbar vattenresursförvaltning och skapas utifrån observerade data. Hydrologiska modeller är nödvändiga för att förutsäga vattenbalansen för tidsperioder och områden utan data, men påverkas av observationsosäkerheter. Metoder för att hantera sådana osäkerheter i vattenresursmodellering är av stor betydelse i regioner såsom Centralamerika.

Observationsosäkerheter hanterades på tre olika sätt i denna avhandling; kvalitetskontroll, kvantitativ uppskattning och utveckling av modellutvärderingsmetoder för beaktande av icke kvantifierbara osäkerheter. Ett viktigt första steg är kvalitetskontroll och samtidig analys av datas representativitet. Vid karaktäriseringen av nederbördsregimen i Cholutecaflodens avrinningsområde i Honduras identifierades fyra olika kvalitetsproblem och 22 % av data sorterades bort. Stationsnätet var otillräckligt för en fullödig karaktärisering av nederbördsregimens variationer i tid och rum. Dessa var mycket stora som ett resultat av komplexiteten hos de nederbördsgenererande mekanismerna.

Kvantitativ uppskattning av observerade datas osäkerhet kan göras när tillräcklig information är tillgänglig. Osäkerheter i vattenföringsdata uppskattades dels vid beräkning av vattenföring med en oskarp regression för en tidsvariabel avbördningskurva, dels från en analys av officiella avbördningskurvor från 35 stationer i Honduras. Osäkerheten var i båda fallen högst vid låga flöden som ett resultat av högre mätosäkerheter samt större naturlig variabilitet än vid höga flöden.

En metod för modellkalibrering med varaktighetskurvor utvecklades och gjorde det möjligt att kalibrera för hela flödesintervallet samtidigt, ta hänsyn till osäkerheter i vattenföringsdata samt kalibrera med icke överlappande driv- och utvärderingsdata. Metoden testades med två olika modeller i två avrinningsområden med olika avrinningsbildningsprocesser, och visade goda resultat jämfört med traditionell modellkalibrering. En *post hoc*-analys gjorde det möjligt att identifiera troliga modellstrukturfel och perioder med disinformativa data. Varaktighetskurvor regionaliserades och användes för kalibrering av en regional vattenbalansmodell för Centralamerika, varvid den initiala modellosäkerheten minskades med 70 %.

Icke representativa nederbördsdata identifierades som det största hindret för regional vattenresursmodellering i Centralamerika. De metoder som utvecklades i detta arbete gör det möjligt att överbrygga ett flertal problem orsakade av bristfällig tillgänglighet och kvalitet av data och leder därmed till en förbättrad uppskattning av osäkerheten i vattenbalanssimuleringar. Sådana osäkerhetsskattningar är ett viktigt underlag vid alla typer av förvaltningsbeslut som rör vattenresurser.

Nyckelord: Avbördningskurva; Centralamerika; GLUE; icke-stationaritet; kvalitetskontroll; modellutvärdering; nederbörd; observationsosäkerheter; avrinningsområden utan vattenföringsdata; oskarp regression; osäkerhetsuppskattning; regionalisering; varaktighetskurva; vattenföring; vattenresurser.

Ida Westerberg, Institutionen för geovetenskaper, LUVAL, Villav. 16, Uppsala universitet, 75236 Uppsala

ISSN 1651-6214
ISBN 978-91-554-8090-5

Till Stina och Gunvor,
som så gärna ville studera men inte
fick samma möjlighet

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Westerberg, I., Walther, A., Guerrero, J-L., Coello, Z., Halldin, S., Xu, C-Y., Chen, D., Lundin, L-C. (2010). Precipitation data in a mountainous catchment in Honduras: quality assessment and spatiotemporal characteristics. *Journal of Theoretical and Applied Climatology*, 101:381-396. © Springer-Verlag 2009, reprinted with permission.

II Westerberg, I., Guerrero, J-L., Seibert, J., Beven, K. J., Halldin, S. (2011). Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras. *Hydrological Processes*, 25: 603–613. © 2010 John Wiley & Sons, Ltd, reprinted with permission.

III Westerberg, I. K., Guerrero, J-L., Younger, P-M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., Xu, C-Y. (2010). Calibration of hydrologic models using flow-duration curves. *Hydrology and Earth System Science Discussions*, 7, 9467-9522. *In review.*

IV Beven, K. J. and Westerberg, I. (2011). On red herrings and real herrings: disinformation and information in hydrological inference. *Hydrological Processes*, 25(10):1676-1680, doi: 10.1002/hyp.7963. © 2011 John Wiley & Sons, Ltd, reprinted with permission.

V Westerberg, I. K., Gong, L., Seibert, J., Beven, K. J., Xu, C-Y., Halldin, S. (2011). Regionalisation of a water-balance model for Central America using flow-duration curves. *Manuscript.*

In Paper **I,** I was responsible for most analyses except the coding of the gap-filling methods and the calculation of the climate indices which were made by A. Walther. The gathering and compilation of precipitation data were made jointly by J-L. Guerrero, Z. Coello and myself, while I was responsible for writing the paper. In Paper **II** I was responsible for writing the paper and performed all analyses, except part of the quality control, which was made by J-L. Guerrero. In Paper **III** I was responsible for writing the paper and did all the modelling and analysis work except setting up and running the TOP-MODEL simulations which was done by P. M. Younger. In Paper **IV** I contributed to the text and arranged the seminar from which the commentary originated. In Paper **V** L. Gong was responsible for the calculation of evaporation, the structure and set-up of the regional model, while I was responsible for the delineation of the catchments and all other analyses in addition to writing the paper. Other co-authors have contributed with ideas, advice and feedback in the work with Paper **I–III** and **V**. Reprints were made with permission from the respective publishers.

In addition I have contributed to the following papers, related to this work but not included in the thesis.

> Kizza, M., Westerberg, I. K., Rodhe, A., Ntale, H. K. (2010). Estimating areal rainfall over Lake Victoria and its basin using ground-based and satellite data. *Submitted to Journal of Hydrology.*

> Guerrero, J-L., Westerberg, I. K., Halldin, S., Xu, C-Y., Lundin, L-C. (2011). Temporal variability in stage-discharge relationships. *Submitted to Journal of Hydrology.*

Contents

Abbreviations

COP	Combined Overlap Percentage
DYNIA	Dynamic Identifiability Analysis
ENSO	El Niño/Southern Oscillation
EP	Evaluation Point
FDC	Flow-Duration Curve
GLUE	Generalised Likelihood Uncertainty Estimation
GRDC	Global Runoff Data Centre
IDW	Inverse-Distance Weighting
OK	Ordinary Kriging
OP	Overlap Percentage
PUB	Predictions in Ungauged Basins
TOPMODEL	TOPographically-based MODEL - Hydrologic catchment model that uses topographic information
UK	Universal Kriging
WASMOD	Water And Snow balance MODelling system – water-balance simulation model

Introduction

Spatial and temporal variability of hydrological processes have a direct influence on many aspects of our lives. In extreme cases, droughts and floods result in loss of life and livelihoods. Less extreme variability may also have large consequences, in economical aspects such as farming, water supply, hydropower, tourism and transportation, but also for natural ecosystems. Knowledge about climate and hydrological processes is the basis for managing water resources and the economy relying on them, for preventing water-related disasters, for understanding human-induced change and for fostering cooperation and avoiding conflict over trans-boundary waters.

The basis for creating such knowledge is the observational data from measurements of hydro-meteorological variables. Observational data are the link between the real-world processes and the perceptual and mathematical models we develop to understand hydrologic behaviour, make predictions for unobserved variability, impacts of human-induced change and future conditions. It is the collision of theory and observed data that generates new knowledge (Box, 1976; Kirchner, 2006). The importance of observational data for understanding these processes has been emphasised by many (Beven, 2002; Box, 1976; Kirchner, 2006; Sivapalan et al., 2003) but also the importance of their intrinsic limitations and uncertainties (Beven, 2002; Kirchner, 2006; Klemes, 1986a; Oreskes et al., 1994; Sivapalan et al., 2003). Observational data are not error free (Eischeid et al., 1995; Pelletier, 1988; Schmidt, 2002; Yang et al., 2006) even if they sometimes may be treated as such. Undetected quality problems can significantly change analysis results, for example the detection of long-term climate trends (Viney and Bates, 2004). Another important source of uncertainty is the difficulty in adequately capturing the often complex space-time variability of hydrological variables with existing monitoring networks (Gottschalk and Askew, 1987).

Remote sensing has increased the availability of many types of data, but the availability of and access to comprehensive observational data is still an important limitation in water-resources modelling in many regions. The number of discharge stations in the world with data reported to the Global Runoff Data Centre database has e.g. decreased substantially in later years from the peak in the late 1970’s of around 4000 stations with daily data to less than 2000 stations in the late 2000’s (GRDC, 2010). This is a result of the degeneration of hydrological monitoring networks as well as decreased and delayed reporting of data. Lack of observational data and degeneration of hydrological monitoring networks is a real problem in many parts of the developing world, especially as human and climate impacts on water resources often can be severe in such regions (Sivapalan et al., 2003). One example is Honduras in Central America where large parts of the hydrological monitoring network were destroyed during the floods that occurred as a result of hurricane Mitch in October 1998, and after which there has been a reduction in both data availability and quality. There are numerous problems related to observational uncertainties in Central America; high spatial and temporal climate variability, and in some countries poor-quality measurement equipment, fragmented and unreliable time series as well as organisational weaknesses such as decentralised responsibility and poor data security (Flambard, 2003). This region is therefore a very "real-world" example for development of methods for dealing with observational uncertainties in water-resources modelling. Lack of appropriate data management such as standardisation, equipment calibration, data security and well-educated staff are important obstacles to efficient water-resources management and hydrological research.

Hydrological monitoring networks have generally been designed for operational purposes and may not be well-suited for scientific purposes (Kirchner, 2006). New measurement techniques are needed, e.g. to characterise subsurface processes and obtain reliable areal data (Beven, 2002; Kirchner, 2006; Klemes, 1986a). The information content and the uncertainties and limitations of presently available data need to be analysed and accounted for when used as a basis for water-resources modelling and management. In addition, there are uncertainties in our perceptual understanding and our models of hydrological systems that affect the reliability of model predictions. Uncertainty in a general sense can be viewed as "an attribute of information" (Zadeh, 2005) or as "a property of the mind" that pertains to an individual (Rougier, 2010). Probability theory has been the traditional way to describe uncertainty (Zadeh, 2005), while other ways include set theory, rough sets, fuzzy sets and Info-Gap theory for extreme uncertainties (Beven, 2009). The nature of uncertainty can be both *aleatory* (a result of the randomness of natural processes) and *epistemic* (non-random uncertainty and a result of incomplete information and understanding). In terms of epistemic uncertainty there are both *known unknowns* and *unknown unknowns*[1] and the latter category is important to remember, especially when it comes to predictions for the future. The methods for dealing with uncertainties depend on their nature and different ways to represent aleatory and epistemic uncertainties in modelling have been suggested (Ferson et al., 2004; Helton and Oberkampf, 2004; Ross et al., 2009).

[1] Donald Rumsfeld, former US Secretary of Defense talked about known and unknown unknowns in a speech on 12 Feburary 2002.

Aim of this thesis

The work in this thesis has been guided by the overall aim to estimate and account for observational uncertainties in water-balance modelling, in order to obtain reliable assessments of available water resources also in basins with scarce or unavailable discharge data. This general aim can be broken down into five specific objectives.

- I Characterise the spatial and temporal variability of the precipitation regime in the Choluteca River basin, Honduras.
- II Calculate discharge and estimate the associated uncertainty based on readily available stage and discharge data.
- III Develop a calibration method that addresses common model-calibration problems of discharge-data uncertainty, sensitivity of performance criteria to flow magnitude, influence of input/output errors of an epistemic nature, and inability to evaluate model performance when observation time periods for discharge and model input data do not overlap.
- IV Develop a method for regionalisation of flow-duration curves to ungauged basins.
- V Assess the potential of calibration to regionalised flow-duration curves in ungauged basins and regional water-balance modelling in Central America.

Observational uncertainties

Precipitation

Characterisation of the space-time variability of precipitation is the starting point for all water-balance studies (Sevruk, 1986). This variability is a result of the complexity of the mechanisms that generate precipitation. A low-intensity frontal rain over a flatland area results in a much lower spatial variability than an intense tropical thunderstorm in a mountainous area. Assessment of areal precipitation in an area with high spatial rainfall variability requires a denser monitoring network than where the variability is low (Michaud and Sorooshian, 1994). In many practical applications the network density is not sufficient, which lead to uncertainties in the characterisation of precipitation variability and as a result uncertainties in discharge simulated with hydrologic models (Brath et al., 2004; Michaud and Sorooshian, 1994; Obled et al., 1994; Wilson et al., 1979). Precipitation has a spatial structure that varies between events, which results in varying over and under-recording by the network. Precipitation uncertainty is many times regarded as the dominant source of observational uncertainty in rainfall-runoff modelling (Gupta et al., 2005).

Indirect precipitation measurement, using radar or remotely sensed satellite data, is generally also affected by large uncertainties, both in the model used to calculate precipitation and the actual measurements (Stephens and Kummerow, 2007; Villarini and Krajewski, 2010). Direct point measurements using manual or automatic bucket-type gauges (the types of data used in this thesis) have the advantage of not relying on an interpretative model, but do not yield spatial estimates. A spatial interpolation method is therefore needed to calculate the distributed precipitation fields or average catchment values used as input to rainfall-runoff models. There is a wealth of methods including nearest-neighbour approaches, inverse-distance weighting and geostatistical approaches like kriging (Isaaks and Srivastava, 1989). The success of the interpolation depends on the strength of the spatial autocorrelation or correlation to a proxy such as topography, and how representative the sampling of the measured data is relative to the underlying spatial variability at the studied time scale, as well as the suitability of the chosen method.

Rain-gauge measurements are affected by a number of error sources. The physical conditions at the location of the gauge (that might change over time, e.g. because of tree growth) result in both random and systematic errors (Sevruk, 1986); random errors can be caused by micro-climatic variations around the gauge while systematic errors relate mainly to wind, wetting and evaporation losses. While some meteorological services correct for systematic errors (Alexandersson, 2003), this is not done in all countries (Sevruk, 1986), and it is sometimes seen as an advantage to provide raw data instead of corrected data. Global and large-scale datasets can demonstrate inhomogeneities because of different procedures. On top of these errors, there are errors of an epistemic type related to malfunctioning automatic gauges and wrongly taken measurements from manual gauges. Viney and Bates (2004) assessed the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high-quality dataset as a result of lack of weekend measurements. They found that 102 out of 181 gauges had hidden untagged accumulations, which, in reanalysing previous studies, led to significant changes in long-term trends at individual stations as well as in rainfall probability and indices of rainfall extremes. Quality assurance of measured precipitation data should be an important first step in any study to remove such errors where detectable.

The importance of precipitation-data uncertainties in a particular model application depends on a number of factors including the magnitude of the errors in the actual point data, the temporal and spatial scale of the precipitation input data relative to the spatial and temporal scale of the precipitation-generating mechanisms, runoff-generation processes and the rainfall-runoff model. Temporal uncertainty can be especially important where only daily-scale accumulations are available while the hydrological processes of interest occur at finer time scales (Kavetski et al., 2011). The effect of spatial precipitation-data uncertainty in hydrologic modelling has been studied relating to monitoring-network density and spatial distribution of rainfall inputs. Except for very large basins (several 10,000 km^2) and catchments with infiltration-excess runoff generation, it has been found that a denser network improves model performance primarily through a better estimate of the total volume of inputs, while the spatial distribution of rainfall is of more importance for the timing of flow peaks (Brath et al., 2004; Lopes, 1996; Michaud and Sorooshian, 1994; Obled et al., 1994; Younger et al., 2009).

Methods for accounting for precipitation-data uncertainty in hydrologic modelling include the use of conditional simulation (Clark and Slater, 2006) and rainfall multipliers (Crawford and Linsley, 1966; Renard et al., 2010). Stochastic perturbation of the inputs through the use of rainfall multipliers in a statistical estimation framework requires specification of the statistical properties of the distribution of the perturbation. This can be difficult even where dense gauge/radar monitoring networks are available (McMillan et al., 2011) and the use of such multipliers has been debated as they will interact with model-structural errors in calibration (Beven, 2009; Yang et al., 2007).

Discharge

Compared to the practical difficulties in measuring precipitation, discharge has often been considered a well-determined hydrological variable. It is an integrated response for the whole watershed and can be determined at a specific location on a river. However, a number of recent studies have shown that there can sometimes be substantial uncertainty in discharge data (Di Baldassarre and Montanari, 2009; Petersen-Overleir et al., 2009), in particular in alluvial rivers with non-stationary river beds (Jalbert et al., 2011; McMillan et al., 2010; Paper **II**). Discharge is traditionally measured indirectly through water stage. A rating curve is fitted to the relationship between stage and discharge which enables calculation of discharge time series from the stage measurements. Rating curves are site-specific and established through repeated stage/discharge measurements at the gauging station.

The stage-discharge relationship can be uncertain for a number of reasons that can be grouped into (1) natural uncertainties, (2) knowledge uncertainty and (3) data uncertainties (Schmidt, 2002). Natural uncertainties include those caused by non-stationary river cross-sections (because of erosion, sedimentation, or other modifications in the channel), growth of vegetation, ice build-up, variable backwater and hysteresis in the stage-discharge relationship during flood-wave propagation (Pelletier, 1988; Reitan and Petersen-Overleir, 2008; Shrestha et al., 2007). Flow that bypasses the gauging structure can also be included in this category. Knowledge uncertainty results from incomplete understanding of the true physical processes and includes improper assumptions in the model of the stage-discharge relationship (Schmidt, 2002). Substantial uncertainty can result from often-needed extrapolation outside the range of the measurements in the rating curve and hysteresis, which is seldom accounted for (Kuczera, 1996; Petersen-Overleir, 2006). There is an overlap between the two categories knowledge and natural uncertainty that Schmidt uses because natural variability results in non-random and non-stationary structure that should be treated as knowledge or epistemic uncertainty – e.g. erosion lowering the channel bed at each major flood or the seasonal effects of vegetation growth on the low-flow rating. Data uncertainties include human-induced observation and data-processing errors (that in part can be addressed by quality control) as well as measurement uncertainties. Stage-measurement errors are usually small whereas the calculation of discharge, normally by integration of the water-velocity field in a river cross-section, can introduce substantial errors (Clarke, 1999; Pelletier, 1988). Measurement uncertainties arise because of insufficient sampling of cross-section geometry and velocity field (both vertically and horizontally) as well as errors in velocity measurement (Pelletier, 1988). Insufficient temporal sampling of stage in calculation of mean daily discharge is an additional source of data uncertainty (Petersen-Overleir et al., 2009).

Discharge uncertainty can affect hydrologic model calibration (Aronica et al., 2006; McMillan et al., 2010) and can be expected to be most uncertain for the highest and lowest flows, which occur most seldom and where the practical measurement difficulties are the greatest. Methods for estimation of discharge uncertainty include both statistical (e.g. Carter, 1970; Di Baldassarre and Montanari, 2009; Herschy, 1970) and non-statistical (e.g. Pappenberger et al., 2006; Shrestha et al., 2007) methods. Traditional statistical methods (such as non-linear least squares) rely on assumptions about the rating errors (homoscedasticity, normal distribution, stationarity, etc.), which if not met can result in biased uncertainty estimates (Petersen-Overleir, 2004; Reitan and Petersen-Overleir, 2008). Bayesian statistical approaches rely on similar assumptions but offer the advantage that prior information about channel properties can be used (Moyeed and Clarke, 2005; Reitan and Petersen-Overleir, 2008). Non-statistical methods are advantageous where the assumptions of the traditional methods are not met, and include those based on single-valued rating curves fitted to randomly chosen subsets of the rated data (Burkham and Dawdy, 1970; McMillan et al., 2010) and fuzzy methods where interdependencies between errors are considered implicitly (Krueger et al., 2010; Pappenberger et al., 2006; Shrestha et al., 2007). Non-stationarity of the cross-section over time as fill, scour and other processes occur in the channel is a problem in many alluvial rivers and violates the assumptions of most methods (Reitan and Petersen-Overleir, 2008; Schmidt, 2002). The traditional approach in calculating discharge under these circumstances is to apply time-variable shifts to the rating curve and/or to make very frequent ratings (Schmidt, 2002). Few studies have quantitatively accounted for rating-curve uncertainty caused by non-stationarity in the stage-discharge relationship (Burkham and Dawdy, 1970; Jalbert et al., 2011; McMillan et al., 2010).

Methods that account for discharge-data uncertainty in model calibration include Bayesian calibration to an estimated probability-density function of discharge (McMillan et al., 2010), Bayesian calibration with a simplified error model (Huard and Mailhot, 2008; Thyer et al., 2009) and limits-of-acceptability calibration in GLUE for rainfall-runoff modelling (Liu et al., 2009) and flood-frequency estimation (Blazkova and Beven, 2009).

Other observational and data uncertainties

Potential evaporation can be calculated indirectly through variables such as temperature, wind speed, relative humidity and net radiation, and is then affected by uncertainties in these data as well as the assumptions of the method used for calculation. The formulation of the Penman-Monteith method used in this thesis (Allen et al., 1998), for example, assumes a uniform grass surface with fixed surface resistance and albedo. Alternatively, pan evaporation measurements can be used, which need to be corrected by empirical pan coefficients to obtain an estimate of potential evaporation. Such measurements depend on the surroundings at the measurement site and how the pan is exposed (Beven, 2001). In most cases errors in precipitation will likely be more important than errors in evaporation (Gupta et al., 2005; Paturel et al., 1995).

Uncertainties in the calculation of physical catchment characteristics can also be important but have in many parts of the world been significantly reduced through remote sensing. High-resolution hydrographic and elevation data are now available for a large part of the globe (Lehner et al., 2008). The precision in the delineation of catchment boundaries depends on the topography with better results in mountainous than in flat areas. From a water-balance perspective the groundwater catchment boundary can be equally important, but is not always possible to infer from surface topography (Skop and Loaiciga, 1998). This is a major difficulty in regions with karst systems where recharge often comes from regions outside the topographic catchment area (Bonacci, 1999). Inter-basin transfers can also have a substantial effect on catchment water-balances in areas with deep volcanic soils (Genereux et al., 2002). Metadata are also affected by uncertainties, e.g. station coordinates that have low precision can result in substantial uncertainty in the station location at a local scale.

Human impacts on the natural hydrological response must also be taken into account; hydrological regimes are affected in a multitude of ways including water-withdrawals, building of dams, and irrigation (Wagener et al., 2010).

Quality control

An important first step in any hydrological study is the quality control and concurrent analysis of the observed data, which serves both to remove detectable errors and explore the characteristics of the dataset. Several studies have stated the need for quality control of hydro-meteorological data in Central America (Aguilar et al., 2005; Balairón Pérez et al., 2004; Flambard, 2003) and other regions (Eischeid et al., 1995; Gonzalez-Rouco et al., 2001; Viney and Bates, 2004).

Errors can relate to malfunctioning measurement equipment as well as human-induced errors like misread and mistyped records and database inconsistencies. The stochastic nature of precipitation, in combination with measurement difficulties, makes quality control of this type of data more problematic than for a spatially continuous variable like temperature or a temporally continuous variable like discharge. Several different approaches have been used for quality control of precipitation data including interpolation and homogenisation (Gonzalez-Rouco et al., 2001), identification of outliers (Eischeid et al., 1995), e.g. using predefined high/low extreme values (Feng et al., 2004), fitting of specific distributions (You et al., 2007) and various statistical and visual analyses for data homogeneity (Feng et al., 2004).

Model evaluation and uncertainty estimation

> ...it is, of course, the *error* signal [...] that can produce learning. The good scientist must have the flexibility and courage to seek out, recognize, and exploit such errors – especially his own.
>
> *Box, 1976*

Modelling of hydrologic processes

The complexities of the hydrological processes that occur in nature can never be fully represented in any model. The catchment is an open system where a multitude of different interrelated energy, vegetation and water processes occur at different temporal and spatial scales. Many of these processes are nonlinear and the boundary conditions are generally poorly known (Beven, 2009). Different hydrologists will understand these processes and their importance in different ways depending on their knowledge and prior experiences. It is therefore important to remember that there is an inherent subjectivity in all modelling, which stems from the background of the modeller and gives her/him a particular pair of "hydrological glasses" when looking at a catchment. As discussed by Rougier (2010), this subjectivity "lies at the very heart of what makes a scientist an expert in his or her field: the capacity to make informed judgements in the presence of uncertainty". The importance of the modeller's personal judgement was manifested in the modeller-comparison study at the artificial Chicken Creek catchment, where predictions (made without calibration to discharge data) from ten modeller groups varied substantially and all failed to represent the observed water balance and hydrological processes (Holländer et al., 2009). The choice of model is for example often highly influenced by the previous experiences of the modeller and the tradition at the university or company where she/he is working. In an aquifer-vulnerability study in Denmark, five different consultancy firms ended up with significantly different model structures and predictions even though they modelled the same site using the same data (Refsgaard et al., 2006). Where, as in this case, there is little knowledge and data to support the selection of an appropriate model structure, the use of different model structures becomes a critical part of the uncertainty estimation.

Many early models were parsimonious lumped conceptual models (see e.g. the review by Clarke, 1973), whereas with the advent of high-power computers more complexity was built into the models. In many cases it was an aim to represent as many processes from the perceptual process understanding as possible at an as fine scale as possible in order to obtain a *realistic* model (exemplified in the blueprint for a physically-based distributed model by Freeze and Harlan, 1969). It was hoped that an extensive representation of small-scale physics would result in better predictions also for changed catchment conditions, ungauged catchments, spatial variability and water quality (Abbott et al., 1986). However, incorporating more processes into the model necessarily leads to more model parameters, and many of these parameters need to be calibrated. The need for model calibration remains as the model is still a very crude approximation of the real-world processes, the descriptive equations are applied for different scales and physical conditions than those for which they were derived and the scale of the model is not commensurate with the scale of the observed data (Beven, 1989; Grayson et al., 1992). The last point is a fundamental problem in catchment hydrology as there are yet no measurement techniques for measuring many water fluxes and storages at the scale of interest or for characterising the complexities of the heterogeneous subsurface (Beven, 2002; Klemes, 1986a). On top of that are the errors introduced by the numerical time-stepping scheme used to solve the model equations (Kavetski and Clark, 2010) as well as limited information content and errors in the observed data (see discussion in Paper **IV**). These approximations and uncertainties involved in any modelling application result in *non-uniqueness* or *equifinality* in representations (model inputs, model structures, model parameter-value sets, model errors) that are consistent with the observed data and therefore uncertainty in model predictions (Beven, 1993; 2009; Oreskes et al., 1994). For hydrologic models calibrated with daily discharge data in snow-free catchments the information content in the data may only be sufficient to identify around four model parameters (Jakeman and Hornberger, 1993).

In modelling the hydrological behaviour of a catchment we would like to get “the right answers for the right reasons” (Kirchner, 2006) but can this be at all possible when “all models are wrong” (Box, 1979)? Strictly speaking we can never be sure if we know the right reasons as model verification is not possible in an open system where we don’t have complete access to the natural phenomena; models as hypotheses can thus only be falsified or confirmed but never verified or validated (Oreskes et al., 1994). Beven (2002) suggests an alternative blueprint in light of the 1969 blueprint of Freeze and Harlan “that explicitly recognises the potential for equifinality in scale-dependent model representations” and in which *physically-based* foremost implies consistency with observations at the scale of interest. At the same time the model should not, as also stated by Box (1976), be “importantly wrong" or over-parameterised. While "all models are wrong", at the same time, "some models are useful" (Box, 1979). The required complexity of the model structure will depend on what types of predictions that are needed for a particular application, e.g. if distributed groundwater-level predictions are needed or if only the discharge at the catchment outlet is of interest. This will limit the number of feasible models and the choice will be further refined in calibration to observed data when models can be rejected as non-behavioural if they are not consistent with the observed data, given their uncertainty. This focus on observational data and its uncertainties was also one of the points made by Kirchner (2006), who emphasises the need for more comprehensive model-evaluation methods that also recognise the intrinsic limitations of the available data. How can we then evaluate whether model predictions are consistent with the observed data or not, given the data limitations? Historically, model-calibration techniques have mostly aimed at finding an optimum or best set of parameters according to some evaluation criterion (Beven, 2009). Such techniques do not account for the uncertainties inherent in the model and data, and the effects of these uncertainties are of critical concern when assessing the reliability of the modelled result – an uncertainty-estimation technique is therefore required.

Uncertainty estimation in hydrologic modelling

All methods for estimation of uncertainty in model simulations and predictions rely, implicitly or explicitly, on a set of assumptions about the nature of the deviations between the simulated and observed data. It is therefore important that these assumptions are stated clearly so that the meaning of the uncertainty in the modelled result can be understood. Such assumptions can be made within different frameworks including Bayesian statistics (Renard et al., 2010; Rougier, 2010), fuzzy sets (Dubois and Prade, 1980) and set-theoretic approaches (Beven, 2009; Keesman and Vanstraten, 1990; Spear and Hornberger, 1980; Whitehead and Young, 1979). There is no general consensus on which method to use (Montanari, 2007) and the choice of a particular method has philosophical undertones. However, the framework of choice can be guided by the (expected) complexity of the total error structure. Where a suitable error model can be identified and confirmed in a posterior analysis (e.g. Engeland et al., 2005; Yang et al., 2007), a Bayesian statistical framework can be useful as it provides the probability of predicting an observation conditional on the model. In Bayesian inference for calibration of a given model M, knowledge about the model parameters Θ_M is formally inferred from the observed data O through Bayes's Theorem:

$$P(\Theta_M|O) = P(\Theta_M) \cdot P(O|\Theta_M)/C \qquad (1)$$

$P(\Theta_M)$ is the prior probability density of the model parameters, $P(O|\Theta_M)$ is the likelihood of simulating the observed data given the model and C is a scaling constant to ensure unity of the posterior probability density $P(\Theta_M|O)$. The likelihood function is derived from the assumptions about the error and in simple cases an additive error model where errors are assumed to be normally distributed and independent is often used. However, the residual errors for hydrologic models are generally not simple, e.g. because input errors are processed nonlinearly through the model, which in itself is not error free (Beven, 2009). More complexity has been included in error models by accounting for autocorrelation, heteroscedasticity and seasonally varying error characteristics (Yang et al., 2007), in which case the uncertainty in the error model was found to dominate the predictive uncertainty.

The structure of the errors may in many cases be even more complex and non-stationary as a result of all the multiple sources of uncertainty discussed earlier, particularly non-stationary unknown input errors. It might then be impossible to find a suitable error model and if an overly simple formal statistical error model is used, this can lead to overestimation of the information content in the data and biased parameter estimates (Beven et al., 2008). In such cases of complex and non-stationary errors that are not easily characterised in an error model, fuzzy set-theory based modelling (Shrestha et al., 2007) or set-theoretic approaches (Beven and Binley, 1992; Hornberger and Spear, 1981; Keesman and Vanstraten, 1990; Wagener et al., 2003) are advantageous as these do not make strong statistical assumptions about the specific structure of the errors and the nature of the uncertainties involved. In set-theoretic approaches like the Generalised Likelihood Uncertainty Estimation (GLUE) method (Beven, 2009; Beven and Binley, 1992) the assumptions are instead made explicit by the choice of model evaluation and rejection criteria, treatment of observational uncertainties in input and output, prior parameter-value ranges, etc. Such methods do not result in predictions that are *probabilistic* but instead *possibilistic* or *non-probabilistic,* meaning that the uncertainty should not be interpreted statistically. This appears to have been misunderstood in some cases, e.g. where results from GLUE used with informal likelihoods were interpreted probabilistically (Stedinger et al., 2008). An approach similar to GLUE is the Dynamic Identifiability Analysis (DYNIA) in which parameter identification is analysed dynamically and periods or high information content for specific parameters can be identified (Wagener et al., 2003).

The explicit or implicit assumptions about the nature of the errors should be tested against calibration-independent data to assess how reliable these assumptions (and therefore the model predictions) are. A strict "validation" is not possible as we know that all models are simplifications of the real-world system they simulate and therefore imperfect; such an analysis might therefore better be termed "evaluation" or "confirmation" (Beven, 2009; Oreskes et al., 1994) and the former term is adopted in this thesis. Evaluation of a calibrated model is normally made through a split-sample test where the model is calibrated on one set of data and then evaluated against another. This is often not a particularly strong test if the model will be used for making predictions outside the range of conditions for which it was calibrated, and Klemes (1986b) suggests differential (test in different climate, land-use, etc conditions), proxy-basin split-sample (test in a different basin in the same region) and differential proxy-basin split-sample (a combination of the two) tests in such cases. Then, a multi-model approach would appear advisable (Refsgaard et al., 2006), especially as many models may be expected to fail differential split-sample tests (Seibert, 2003). Apart from visually inspecting the correspondence between the simulated time series and the observed data in calibration and evaluation, a number of post-calibration diagnostic tests can be made. Rank histograms or predictive quantile-quantile plots (Thyer et al., 2009) analyse the quantiles of the observed values in the simulated distribution but do not account for uncertainty in the observed data, as does the generalised rank histogram (McMillan et al., 2010). Scaled scores to limits of acceptability is a further extension, as the relative deviations from the data-uncertainty limits for the predictions that are outside the limits are also considered (Paper **III**).

Observational uncertainties and model evaluation

The observational data available for modelling the hydrologic behaviour of a catchment vary greatly in quality and quantity; often there might only be discharge and climate data at hand (in the best case these are continuous series and measured in the catchment of study). Other types of model-evaluation data such as groundwater levels and isotopes are not as frequently available. The goal of model calibration should be to extract the available information from the observational data while at the same time recognising the intrinsic limitations of the data (see e.g. the discussion by Kirchner, 2006), which results in a balancing of Type I (false positive) and Type II (false negative) errors within the constraints of the observational uncertainties (Beven, 2010; Paper **IV**). This can be difficult to achieve in practice as the information content is generally unknown as a result of unknown errors in the data, e.g. because of non-representative sampling of the measured processes such as when rain cells pass in-between rain gauges. Information content in the context of model calibration has been analysed in terms of how many discharge data points are needed to calibrate a hydrological model and how such points should best be spaced in time (Juston et al., 2009; Seibert and Beven, 2009). Measurements during a recession period, where discharge is highly autocorrelated in time, will for example provide information about both absolute discharge and rates of change (Seibert and Beven, 2009). If measurements are taken at hydrologically-informed times, a small fraction of the data may contain almost all of the information content (Juston et al., 2009). If a poor (random) selection of observation days is made this can actually worsen the results as a result of model-structural error or uncertain or disinformative data (Seibert and Beven, 2009).

In a broad sense, two ways of defining how consistent model simulations are compared to observational data can be distinguished from the literature; either the likelihood (formal or informal) is based on the error series (e.g. Choi and Beven, 2007; Nash and Sutcliffe, 1970; Thyer et al., 2009) or it is based on how well simulations reproduce information, such as a flow-duration or recession curve, calculated from the data (Blazkova and Beven, 2009; Gupta et al., 2008; Montanari and Toth, 2007). When the first approach is used, analysis of information (used for calibration in the second approach) has been utilised as a posterior analysis of the simulated results (Houghton-Carr, 1999; Kavetski et al., 2011; Son and Sivapalan, 2007). The reverse, a posterior analysis of the error series after calibration to the information, was used in this thesis as a way to learn about probable model-structural and data errors (Paper **III**).

Traditional informal model-evaluation criteria like the much-used Nash-Sutcliffe efficiency are included in the first category (Nash and Sutcliffe, 1970). It is a normalisation of the mean squared error by the variance of the observed data and varies between minus infinity and 1.0. Despite its widespread use, the suitability of this criterion has been much debated (Criss and Winston, 2008; Gupta et al., 2009; Schaefli and Gupta, 2007; Seibert, 2001). Other criteria have been proposed that focus on other aspects of the hydrograph (e.g. Krause et al., 2005) as well as the use of several criteria together in a multi-criteria calibration (Boyle et al., 2000; Gupta et al., 1998) to constrain different aspects of the hydrograph simultaneously. Problems with these types of lumped evaluation criteria in uncertainty estimation include: evaluation-data uncertainty is generally not accounted for, there are no thresholds of what is acceptable in terms of the criteria values, and the exact weighting of the performance at different parts of the hydrograph depends on the hydrograph characteristics. Other approaches based on the error series do account for uncertainty in evaluation data, for example Bayesian calibration to discharge uncertainty derived from rating-curve analyses (Huard and Mailhot, 2008; McMillan et al., 2010; Thyer et al., 2009). Another way of accounting for observational uncertainties is by using limits of acceptability in the extended GLUE framework (Beven, 2006). These limits should be defined prior to running the model and specify the estimated uncertainty in the evaluation data. In calibration, all simulations that are within these limits are considered behavioural. Liu et al. (2009) apply such limits to time series of discharge where the discharge-data uncertainty was derived from the rating curve and find no simulations that were completely consistent with the limits.

The second category, using hydrologic "signatures" or information derived from discharge (or other types of data), is usually multi-objective as a single constraint is not sufficient to constrain the model simulations. Many of these studies have been made within a set-theoretic approach for uncertainty estimation (e.g. Blazkova and Beven, 2009; Winsemius et al., 2009; Yadav et al., 2007), but Bayesian statistical approaches have also been used (Bulygina et al., 2009; 2011). Few studies have explicitly incorporated an estimation of the observational-data uncertainty in the analysis (e.g. Blazkova and Beven, 2009). The types of information that have been used include recession curves (Winsemius et al., 2009), slope of the flow-duration curve (Yadav et al., 2007; Yilmaz et al., 2008), base-flow index (Bulygina et al., 2009), spectral properties (Montanari and Toth, 2007), and snow-water equivalent (Blazkova and Beven, 2009). Another approach is to define evaluation criteria based on exceedance percentages of the flow-duration curve (Refsgaard and Knudsen, 1996; Yu and Yang, 2000).

In terms of uncertainty estimation, this latter approach may be seen as resulting in more directly interpretable uncertainty bounds as they relate directly to the hydrological information and its estimated uncertainty. Another advantage with the methods in this latter category is that such information can be useful in situations with data scarcity (Montanari and Toth, 2007; Winsemius et al., 2009), and for PUB through regionalisation of the hydrologic information, which is then used to constrain the model parameters for the ungauged-basin simulation (Wagener and Montanari, 2011; Yadav et al., 2007). The spatial coverage of discharge stations is not sufficient in many of the world's basins and many basins are ungauged or poorly gauged. There are no complete databases of the world's discharge stations but some studies report that approximately 50% of the continental land mass is gauged (Fekete et al., 2002). This might then include large basins that are insufficiently gauged at a local scale. The lack of discharge stations makes PUB an important prerequisite for a comprehensive mapping of the spatial and temporal variability of water resources (Sivapalan et al., 2003). Conceptual water-balance models have traditionally been regionalised by transferring optimally-calibrated parameter values from gauged to ungauged basins using some measure of hydrologic similarity or a regression with physical characteristics of the basins (Parajka et al., 2005; Seibert, 1999; Vandewiele and Elias, 1995; Xu, 2003). Such procedures are limited by their assumption of model-parameter independence and incomplete assessment of predictive uncertainty for gauged and ungauged basins (Buytaert and Beven, 2009; McIntyre et al., 2005; Reichl et al., 2009). Ensemble predictions using model averaging, based on combined likelihoods from model performance in gauged basins and hydrologic similarity with ungauged basins, have been used to address these limitations (McIntyre et al., 2005; Reichl et al., 2009). This approach allows for treatment of uncertainties in model structure, basin attributes, and input and output data. Studies using regionalised information include those by Yadav et al. (2007) who regionalise constraints on expected watershed response behaviour in the UK, and Bulygina et al. (2009) who regionalise the base-flow index. Yu and Yang (2000) regionalise flow-duration curves and calibrate their model against a performance measure based on specific exceedance percentages of the FDC but do not estimate predictive uncertainty.

Study areas and data

The studies presented in this thesis were all conducted with data from the Central-American region except the study in Paper **III** where data from a British catchment were used in addition to those from a Honduran basin. The British catchment is described in Paper **III**, but here the focus is on the Central-American region.

Central America is a region with a highly variable climate in both space and time, despite its small areal extent (around 520,000 km^2). One reason for this is the high mountain range that stretches through the region and reaches maximum elevations of 4,200 metres in Guatemala (*Figure 1*). The rivers that drain to the Atlantic are generally longer and with larger draining basins than those that drain to the Pacific. The largest lake in the region, Lake Nicaragua (approx 8,000 km^2), is the second largest in the whole of South America. Around 40 million people live in Central America with most of the population on the Pacific side of the isthmus. Agriculture is an important part of the Central-American economies with crops such as coffee, banana and sugarcane. Deforestation and the loss of fertile soils is a large problem in the region and it is enhanced by the steep topography in many areas.

The high spatial and temporal variability in climate results in many water-related disasters such as droughts and floods of which the hurricane Mitch in October 1998 was the most disastrous in modern time. The consequences of flooding have been severe in many Central-American countries: inundation and destruction of important crops, promulgation of landslides, loss of life and private property as well as destruction of public infrastructure (Waylen and Laporte, 1999). Sustained droughts can have severe consequences for hydro-power generation, water supply and irrigation and also lead to loss of tourism, which is an important economic sector in many parts of the region (George et al., 1998). Heavy reliance on hydro-electric power and limited transport infrastructure mean that these countries are particularly vulnerable to negative economic consequences of flooding (George et al., 1998; Waylen and Laporte, 1999).

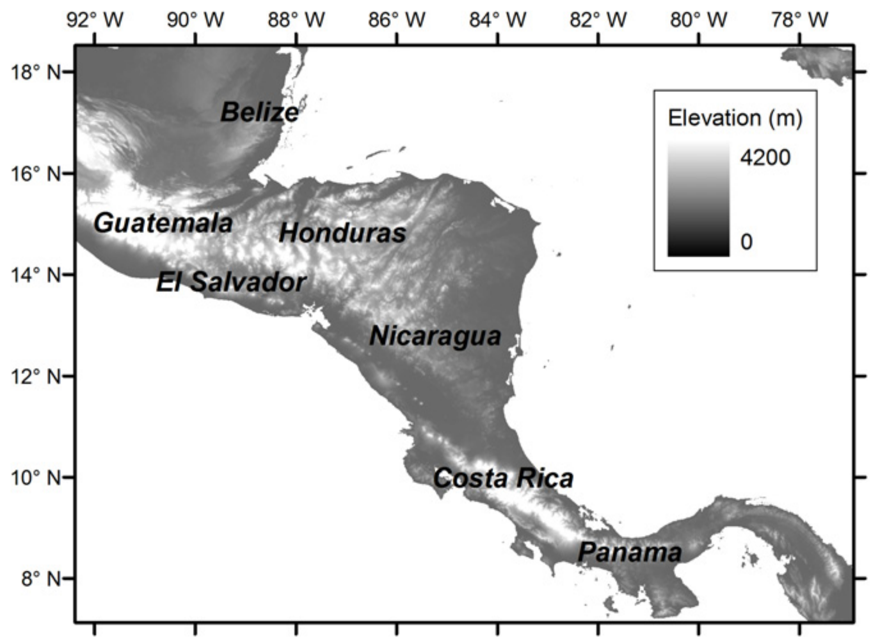

Figure 1. The Central-American region and the elevation distribution (in metres).

The characteristics of the complex regional climate have been studied extensively (Alfaro, 2002; Amador, 2003; Amador et al., 2006; Enfield and Alfaro, 1999; Magaña et al., 1999) but there are few hydrological studies in the published literature (George et al., 1998; Maurer et al., 2009; Waylen and Laporte, 1999). One reason for the scarcity of peer-reviewed literature is the difficulty to access comprehensive and good-quality hydrometeorological data. The importance of quality control of observational data is stated by several studies (Aguilar et al., 2005; Flambard, 2003; Paper **I**). The precipitation regime has a less marked seasonal variability on the Atlantic Coast compared to the Pacific Coast where around 80% of the precipitation falls in the rainy season from May to the end of October (Portig, 1976). On the Pacific Coast there is a rainfall minimum, the so-called midsummer drought or *veranillo* in July–August, creating a bimodal regime with rainfall peaks in June and September–October (Magaña et al., 1999). The spatiotemporal variability of precipitation is high since precipitation is mainly convective and generated by mechanisms such as hurricanes, tropical storms and depressions, and easterly waves in the atmosphere (Peña and Douglas, 2002). Temperature variability is low, with a greater diurnal than annual range, which is characteristic of the tropics. Climate variability on an interannual time scale is pronounced with large differences between wet and dry years; this variability is modulated by ENSO (El Niño/Southern Oscillation) and Atlantic sea-surface temperatures (Diaz et al., 2001; Enfield and Alfaro, 1999).

Methods

Quality control of observational data

A combination of automated tests and visual data inspection was used for quality control of precipitation, stage and discharge data in Paper **I** and **II**. In Paper **V** a visual analysis was used to remove obvious outliers in discharge data and assess inhomogeneities in precipitation data. The quality control of precipitation data in Paper **I** was focused on four types of quality problems detected in the data: (1) obvious outliers were first removed and data were then flagged for three types of homogeneity errors recurrent in the dataset; (2) too-frequently occurring data; (3) sequences of too-low data, and (4) dry months in the rainy season. All data series were inspected visually on a daily and monthly level, and flagged data were removed if deemed necessary. This process was subjective in the cases where errors were not obvious, but supported by double-mass curves, scatter plots and correlation analysis with surrounding stations of good quality. The quality control of stage and discharge data in Paper **II** was mainly performed visually; outliers in the stage-discharge data could many times be readily identified by comparison to the time series of stage measurements.

Characterisation of precipitation variability

Rain-gauge data were used to characterise the precipitation variability in the Choluteca River basin through calculation of climate indices and spatial interpolation of the quality-controlled gauge data (Paper **I**). The climate indices characterised the intensity and seasonal variability of the precipitation regime, e.g. through the number of heavy precipitation days and the start and end date of the rainy season. Shorter gaps in the data were filled before spatial interpolation was performed at a monthly and mean annual scale using three methods of varying complexity (inverse-distance weighting, IDW, ordinary kriging, OK and universal kriging, UK). Two different data periods (1985–95 and 1996–05) were used to identify differences in the results relating to varying spatial station density. Spatial interpolation methods are based on assumptions about the pattern of spatial continuity (Isaaks and Srivastava, 1989). In IDW it is assumed that distance alone explains the weight of the different stations in the interpolation at unmeasured locations. In kriging a stationary random function model (estimated from the data in the form of a sample semi-variogram) describes the pattern of spatial continuity. For UK a trend is modelled as a function of the coordinates and subtracted where after the residual semi-variogram is calculated.

Estimation of discharge-data uncertainty

Discharge-data uncertainty was calculated in two ways; through the use of a time-variable rating curve in combination with time series of stage data for the Paso La Ceiba station in Honduras (Paper **II**) and for 35 stations in Honduras using official rating-curve equations and associated stage-discharge data (Paper **V**).

The stage-discharge relationship at the Paso La Ceiba station on the alluvial Choluteca River is non-stationary as a result of erosion, deposition and other processes taking place in the river channel. There have been few studies that describe and quantify the uncertainty pertaining to discharge calculated from non-stationary rating curves, even if such non-stationarity could be expected to occur in many alluvial rivers (Jalbert et al., 2011; McMillan et al., 2010). There were over 1,200 ratings available in 1980–1997, and we used a weighted fuzzy regression of the rating data based on estimated uncertainties in the stage (water depth) and discharge measurements. The regression was performed within a moving time window of 30 data points at a time. The discharge data were first Box-Cox transformed and the stage data were log-transformed to obtain a linear relation and there was considerable spread in the ratings, especially in the low-flow range. Uncertainties in stage and discharge were estimated as constant percentage errors at ±5% and ±25%, and used to define the triangular fuzzy numbers used in the regression. The fuzzy regression followed Hojati et al. (2005), but was modified so that the regression was solved by minimising the deviations between the predicted and observed uncertain intervals for the upper left and the lower right corners of the fuzzy stage-discharge representation (d_{ilU} and d_{irL} in *Figure 2c*). Each data point in the moving window was weighted, with the largest weight given to the most recent measurement to account for non-stationarity in gauge height within the moving window (Paper **II**). The three highest discharges were included in all windows to constrain the upper range when high flows were missing, but were given small weights. There were only three measurements of gauge height per day (06:00, 12:00 and 18:00 h) while the flow peaks were often of short duration (~6–12 h). This can result in a temporal commensurability error in the calculation of mean daily discharge, and this error was estimated using 15-min resolution stage data that were available for a later period.

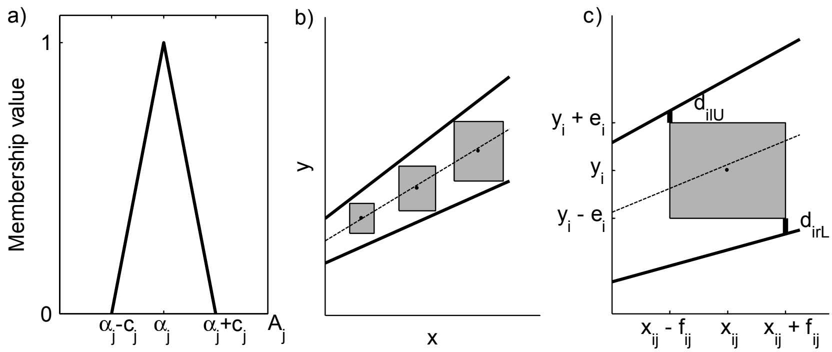

Figure 2. a) Triangular fuzzy regression coefficient, b) Fuzzy linear relationship and c) Variables in the linear-programming solution to the fuzzy regression. The dots in b) represent the crisp observed data and the grey area the fuzzy representation of the uncertainty in the observed data. The upper and lower thick lines represent the limits, between which the membership value of the predicted variable is positive. The dashed line represents the regression estimate with a membership value of one. In c) y_i is the midpoint of the fuzzy discharge and x_{ij} the midpoint of the fuzzy gauge height, e_i and f_{ij} are the half-widths of the fuzzy discharge and gauge height, respectively, and d_{ilU} and d_{irL} are the respective absolute deviations to the upper and lower regression lines. Modified from Hojati et al. (2005).

In Paper **V**, officially calculated discharge data were used as there was no scope for calculating discharge and the associated uncertainty for all the stations in the regional model. Lack of metadata about the characteristics of the gauging stations, gauged data and discharge-data quality is a common problem in hydrological modelling. Here we addressed this problem by making a general estimation of discharge uncertainty using stage-discharge data and corresponding official rating curves for 35 stations in Honduras in an analysis of rating-curve residuals, calculated as a percentage of discharge normalised by mean discharge. This estimation was then used for the whole region, assuming similar monitoring procedures as in Honduras.

Model calibration using flow-duration curves

A method for calibrating hydrologic models using flow-duration curves (FDCs) was developed in Paper **III**. The FDC describes the relation between the magnitude and frequency of stream flows for a particular period of record, but an annual interpretation can also be made (Vogel and Fennessey, 1994). The method was developed within the GLUE limits of acceptability framework (Beven, 2006; 2009) where limits of acceptability are set based on estimated evaluation-data uncertainty. The method was evaluated in the Paso La Ceiba catchment (also used in Paper **II**) and the Brue catchment in the UK, with discharge simulated by the WASMOD model (Xu, 2002) in the first case and dynamic TOPMODEL (Beven and Freer, 2001) in the second. Monte Carlo simulations with 100,000 and 50,000 runs were made with parameter values sampled uniformly from predefined intervals. A number of evaluation points (EPs) were chosen where the FDC of the simulated discharge was compared to the observed FDC. Simulations were defined as behavioural if the simulated FDC was inside the limits of acceptability for the observed FDC at all EPs. The choice of these EPs can be made in different ways depending on the aim of the study and the characteristics of the hydrograph. The high-flow part of the FDC, which describes the dynamic response of the catchment to the effective precipitation input, usually contains most of the information about catchment response and many parameters are therefore sensitive with respect to these high flows. Sufficient points on this part of the FDC were therefore needed and we explored two methods for EP selection. In the first, EPs were chosen based on equal intervals of discharge in-between the minimum and maximum discharge. In the second, the area under the FDC (which represents a volume of water contributed by flows smaller than a certain magnitude) was divided into equal intervals. Performance measures based on a triangular weighting around the best-estimate discharge ($R_{FDC\text{-}Q}$ and $R_{FDC\text{-}V}$ for discharge and volume intervals respectively) were used to calculate informal likelihoods for each simulation (*Figure 3*). The FDC-calibration was compared to calibrations using the Nash-Sutcliffe efficiency (R_{eff}) for different behavioural threshold values in split-sample tests. The performance was evaluated in a posterior analysis in terms of the overlap between the simulated and observed uncertain discharge ranges. The overlap was calculated as the percentage of time steps where the ranges overlapped (OP) and as a combined overlap percentage (COP) which was calculated as the mean of the overlapping interval as a percentage of the observed and the simulated ranges at a specific time step.

$$COP = mean\left(\frac{QR_{overlap}}{QR_{obs}}, \frac{QR_{overlap}}{QR_{sim}}\right) \qquad (2)$$

$QR_{overlap}$ is the intersection between the simulated and observed discharge ranges, QR_{obs} the observed discharge range and QR_{sim} the simulated discharge range. This number was then averaged for all time steps. Scaled scores (*Figure 3a*), which describe the deviation from the observed discharge range at a specific time step, were also used to analyse the simulations.

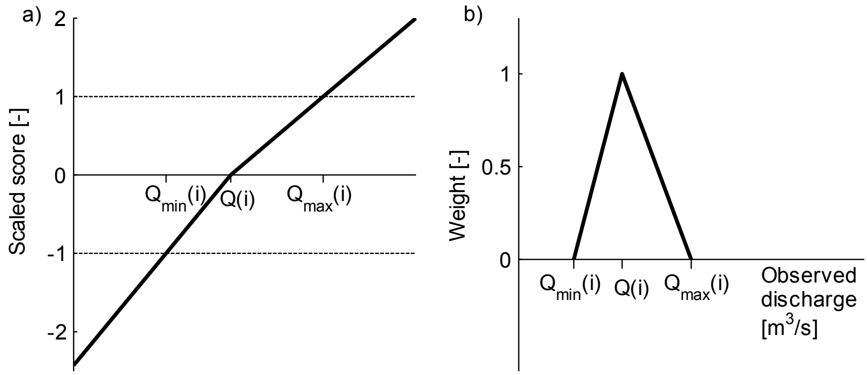

Figure 3. a) Calculation of the scaled scores, $Q_{min}(i)$ is the lower limit for the discharge uncertainty at the *i:th* evaluation point (EP) or time step, $Q_{max}(i)$ the upper limit and $Q(i)$ the crisp discharge. A simulated value that is at the crisp value gets a scaled score of 0, if the value is at the lower limit a scaled score of -1 and at the upper limit of 1, values within or outside are linearly inter- or extrapolated; b) Triangular weighting function applied at each EP such that weights are zero for scaled scores outside the range [-1, 1].

Regionalisation of flow-duration curves

Two methods for regionalisation of FDCs to ungauged catchments were tested in Paper **V**. The FDC uncertainty for each catchment used in the regionalisation was first estimated from discharge-data uncertainty as well as realisation uncertainty if the number of years with data used to calculate the FDC was shorter than the modelling period. The uncertain discharge at each EP on the FDC was defined as a fuzzy number with a triangular membership function defined by the lower, crisp (best-estimate) and upper uncertainty limits. Both methods were based on a weighted linear combination of the FDCs for the hydrologically most similar catchments, using a method similar to that of Holmes et al. (2002). Hydrologic similarity was calculated as a Euclidean distance in the space spanned by standardised catchment characteristics (Burn, 1990a; b; Holmes et al., 2002). Descriptors that characterised the seasonal precipitation variability (e.g. standard deviation of daily precipitation) were used in addition to physical catchment characteristics (e.g. elevation range), as the precipitation regime is highly variable throughout Central America and has a direct influence on the hydrological regime and hence the FDC (Alfaro, 2002; Waylen and Laporte, 1999).

In the first method (R1), a general weighted mean operator for fuzzy numbers (Dubois and Prade, 1980) was used to aggregate the *N* individual fuzzy discharges to a regionalised estimate (*Figure 4*). The individual membership functions were weighted by the inverse of their hydrologically similarity (as expressed by the Euclidean distance) with the target catchment.

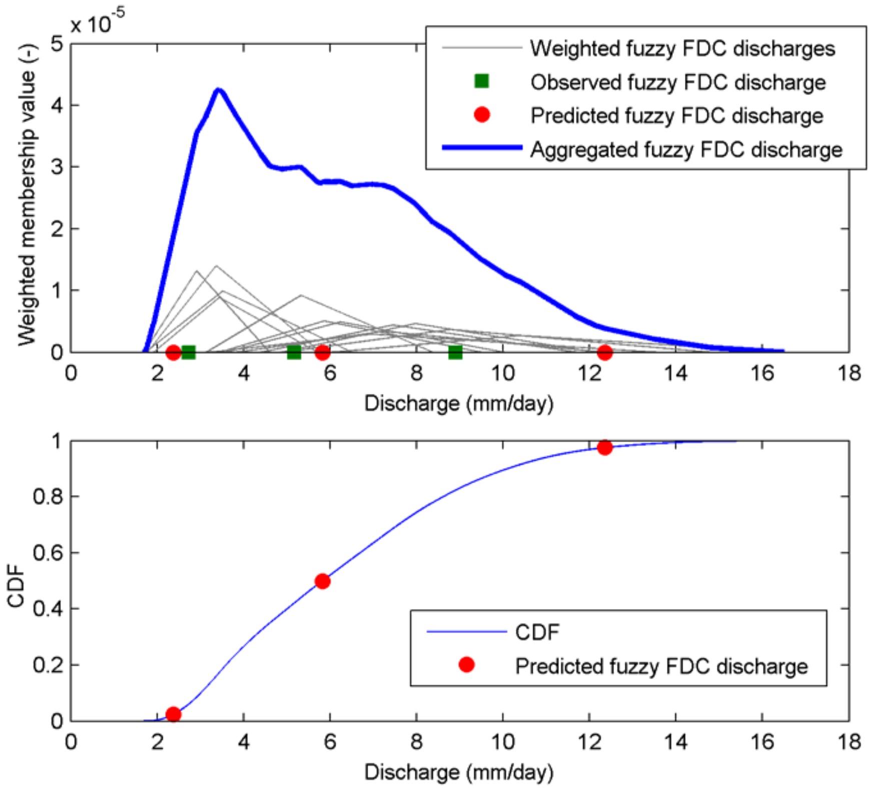

Figure 4. Regionalisation of the uncertain FDCs using the R1 method.

The 2.5, 50 and 97.5 percentiles of the cumulative distribution of the aggregated membership function were used as the regionalised lower, crisp and upper uncertainty bounds. The second method (R2) was a simple weighted linear combination of the upper, lower and best-estimate FDC discharges at each catchment using the *N* most similar catchments. The FDC-regionalisation was evaluated with a jack-knife cross-validation by excluding one catchment at a time.

A regional water-balance model for Central America

In Paper **V**, a regional water-balance model was set up for Central America and calibrated using the regionalised FDCs in combination with the FDC-calibration method developed in Paper **III**. The input data were gridded daily precipitation data from the CRN073 dataset and potential evaporation calculated from variables from the WATCH Forcing Data (Weedon et al., 2010), while discharge data from the Global Runoff Data Centre (GRDC, 2010) were used for evaluation and regionalisation. The lumped conceptual four-parameter WASMOD model (the same as in Paper **III**) was run for the period 1965–94 when the best input and output data were available in terms of quantity and quality. The HydroSHEDS elevation data were used to calculate upstream areas and delineate catchments (Gong et al., 2011). Catchments with major dams were excluded, and 36 non-nested catchments that had sufficient discharge records in terms of quality and quantity of data were selected (*Figure 5*).

The run-off ratios (long-term ratio of run-off over precipitation) were calculated for all catchments and the Budyko curve (Budyko, 1974), which shows the relationship between one minus run-off ratio and the aridity index (potential evaporation divided by precipitation), was analysed to assess the range of climatic conditions in the region and identify stations with unreasonable data (*Figure 2* in Paper **V**). Four catchments had unreasonable run-off ratios (>>1) and were excluded, leaving a final 32 catchments for the regionalisation. The four excluded catchments were all smaller catchments in the mountainous parts of Costa Rica (maximum elevations 1,800–3,000 m a.s.l.) and the precipitation data at a scale of 0.5° were likely not sufficiently representative for these catchments. Two catchments (Laja Blanca and Boca de Cupe in Panama) stood out as having a combination of aridity index and run-off ratio that deviated from the rest but were kept. This appeared to be a result of overestimated precipitation in the CRN073 data set as compared to local Panamanian data.

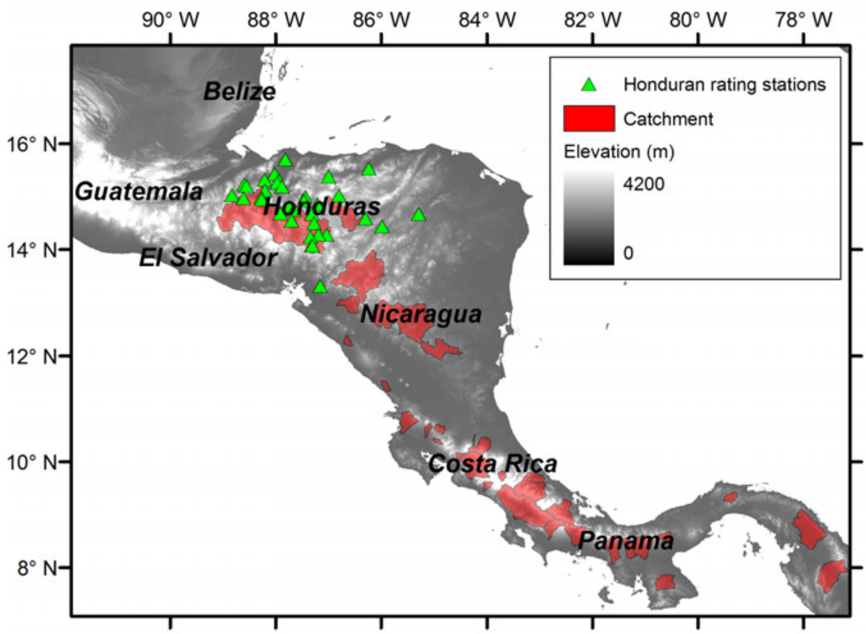

Figure 5. The Central American region, elevation distribution (in metres) and the location of the catchments and Honduran rating stations.

Monte Carlo simulations with 100,000 model runs were made for all catchments and the model was first calibrated using all available discharge data at each station in 1965–1994 for both local and regional EPs. The results were evaluated in a posterior analysis in terms of the overlap between the simulated and observed discharge intervals for low, intermediate and high flows respectively.

Results

Quality control of observational data

In total 22% of the daily precipitation data from the 60 stations in the Choluteca River basin were deleted in the quality control and the percentage of poor data was relatively constant for the whole period 1970–2005 (Paper **I**). The rejected data consisted of 3% erroneous zeros, 11% too-low values, 42% too-frequent values whereas the remaining 44% had other problems mainly relating to homogeneity. Many of the quality problems were related to wrongly taken (even faked) measurements and data-digitisation errors (*Figure 6*). For the discharge data in Paper **II** 1,216 out of the original 1,268 ratings remained after the quality control. The types of identified errors included, e.g., unrealistically low stage values in the midst of recession periods for which obvious cases could be corrected. Visual inspection of the time series of daily precipitation data in the regional data set used in Paper **V** revealed in-homogeneities that were likely related to varying station density and the use of malfunctioning automatic gauges. The modelling time period was therefore chosen to exclude as much of these data as possible.

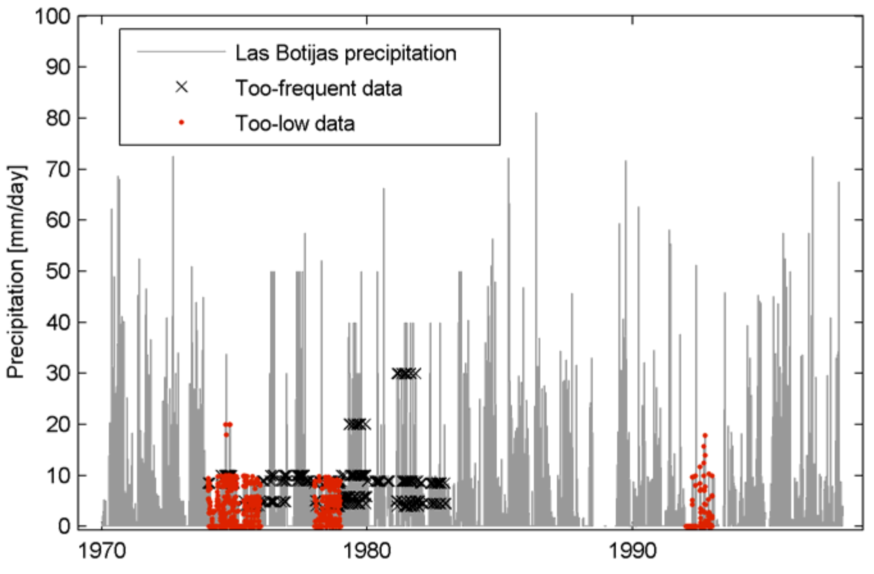

Figure 6. Results of quality control of daily precipitation data from the Las Botijas station; the too-low and too-frequent data were flagged as potentially erroneous.

Characterisation of precipitation variability

The climate indices manifested the large variability in the precipitation regime in the 7,500 km^2 Choluteca River basin as well as the greater intra-annual variability in the southern part close to the Pacific coast compared to the northern part. Because of the high spatial variability, the available gauges did not have a sufficient spatial coverage and large differences were seen between the earlier and later periods (*Figure 7*). This was not believed to be a result of climate variability. The difference between the two periods was greater than between the three interpolation methods. With an increased spatial sampling it is likely that more variability would have been revealed, especially in the upper mountainous parts of the basin. Time series of (average) areal precipitation used as input to hydrologic models will be affected by time-varying errors as a result of the varying station density in combination with insufficient sampling of localised precipitation events, especially at a daily or finer time scale.

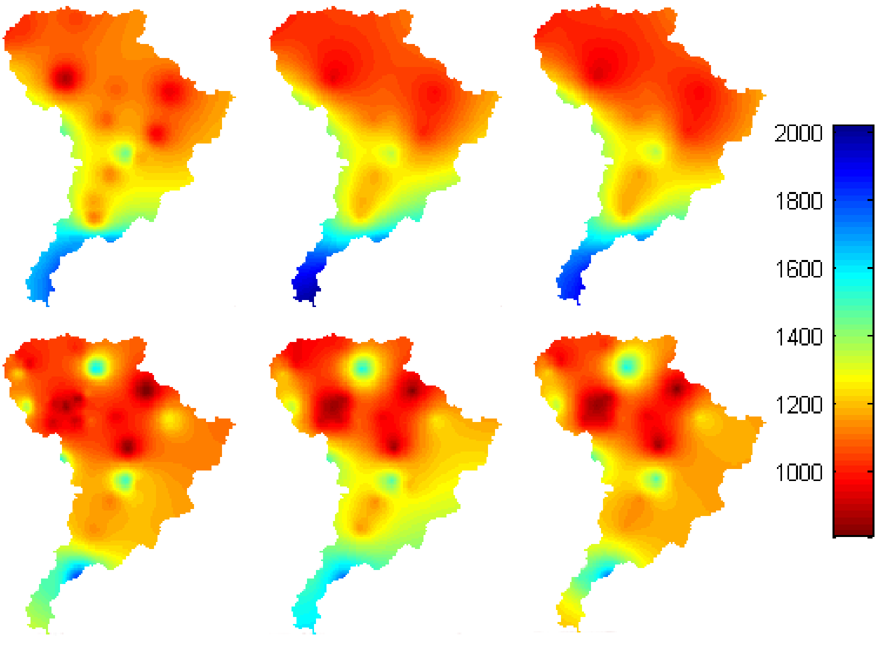

Figure 7. Upper panel: mean annual precipitation (mm) 1975–85 (for the 28 stations with more than 50% complete years with monthly data during the period). Lower panel: mean annual precipitation (mm) 1990–2005 (for the 34 stations with more than 50% complete years with monthly data during the period). The precipitation data were interpolated with inverse-distance weighting (left), universal kriging (middle) and ordinary kriging (right).

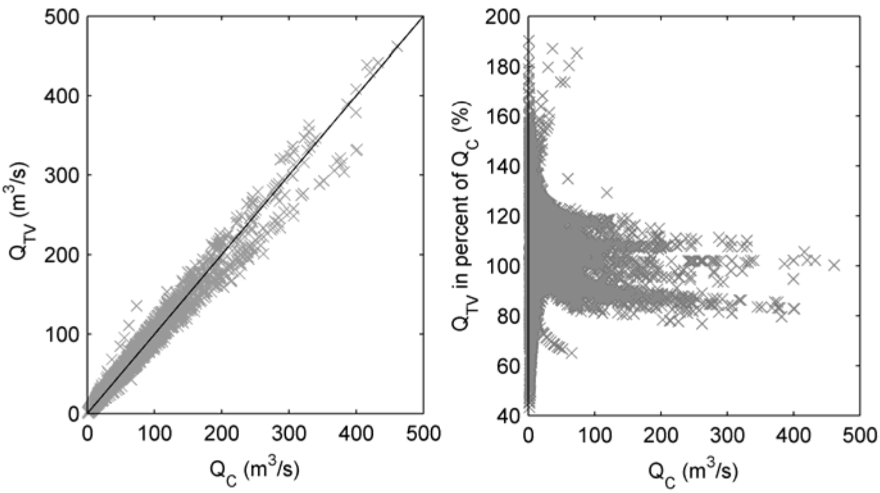

Figure 8. Crisp discharge at a sub-daily time scale calculated with the time-variable rating curve (Q_{TV}) versus discharge calculated with a constant rating curve (Q_C, left). Q_{TV} in percentage of Q_C plotted against Q_C (right).

Estimation of discharge-data uncertainty

The rating-curve parameters varied considerably over time and there were both gradual and abrupt shifts in the stage-discharge relation. Hypothetical discharge that was calculated for two constant gauge heights of 0.5 m (low flow) and 1.5 m (medium-range flow) varied substantially over time with greater variability in the 1980s compared to the 1990s. The variability was particularly large around 1983 when only low-flow measurements were available. In general the largest variability was found during low-flow periods when the lower part of the curve changed shape and this was likely a combined effect of channel changes and measurement uncertainty (that can be high for very low water velocities, Pelletier, 1988). The effect of the time-variable rating curve was assessed by comparing the crisp discharge calculated with this method to the discharge calculated with a constant, single-valued rating curve fitted by linear regression to all the transformed rating data (*Figure 8*). The difference between the two datasets was in the range ±20% for intermediate- and high-range flows while there were greater differences for low flows (-60 to +90%).

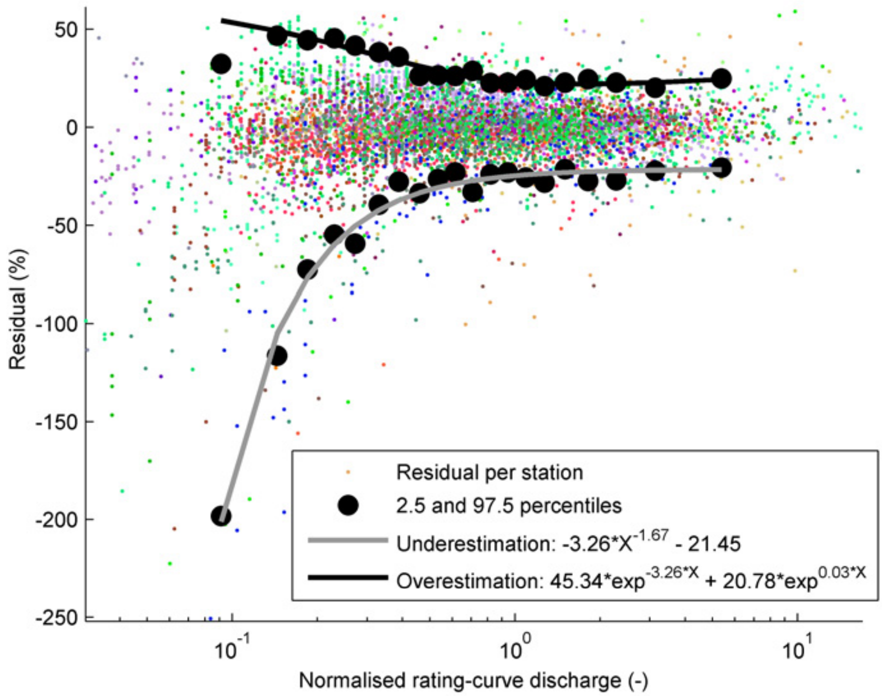

Figure 9. Rating-curve residuals for 35 Honduran stations (one colour per station) and 2.5 and 97.5 percentiles of the residuals in each group plotted against the median normalised discharge in each group. The groups were binned by frequencies of 1, 5, 10, ...95, 100%. The residuals were calculated as rating-curve discharge minus observed discharge (as a percentage of the rating-curve discharge) and the plot excludes a few smaller and larger residuals to improve the visibility for the main flow range.

A similar behaviour, with larger relative uncertainties for low-flows than for high flows was seen in the analysis of the rating-curve residuals for the 35 Honduran rating stations (*Figure 9*) in Paper **V**. The main difference to the Paso La Ceiba station was that the low-flows were underestimated at some stations as a result of poorly fitted rating curves to the lower end. At some stations there appeared to be trends in the residuals for the whole range of flows as a result of poor rating-curve fits. In the high-flow range, most stations had residuals in the range of ±25% of the rating-curve discharge. The estimate of ±25% measurement uncertainty in discharge in Paper **II** therefore appears reasonable, even if the use of a constant percentage uncertainty is only approximate as higher uncertainties could be expected for very low flows.

Model calibration using flow-duration curves

For both models that were used to test the FDC calibration, the R_{FDC-V} calibration better constrained the parameters that controlled the slow-response, evaporation and recession behaviour than R_{FDC-Q} and R_{eff}. This was also manifested in the simulated FDCs and an expected result as no low-flow EPs were used for R_{FDC-Q}, and for R_{eff} the largest weight is put on high and intermediate flows. The posterior analysis showed that the overlap between the simulated and observed discharge intervals was better for the R_{FDC-V} in calibration and evaluation and that the simulated discharges were more accurate for Paso La Ceiba than for Brue. An analysis of scaled scores for different parts of the hydrograph (high flows, low flows, troughs, rising limbs and falling limbs) showed that these distributions were generally more centred on the best-estimate discharge for R_{FDC-V}. The largest differences between the different performance measures occurred for low flows. However, peaks and rising limbs for TOPMODEL (that was evaluated at an hourly scale) showed more underestimation as a result of greater uncertainty in the timing of flow peaks (examples of peak-flow timing are seen in *Figure 16* in Paper **III**). These results suggest that additional criteria will be needed in cases where the timing of peak flows is of great importance, but for water-balance studies this does not appear problematic. This effect was not observed at the daily time step, and (depending on the flow regime) is most likely important for models running on sub-daily time scales. For TOPMODEL, a period of probable model-structural error (in July–Nov 1997) could be detected in the low-flow scaled scores. These corresponded to a period of consistent over-prediction by the model for all performance measures. The WASMOD simulations at Paso La Ceiba revealed the presence of several obviously disinformative events in the modelling data such as lack of observed discharge peaks for heavy precipitation events (*Figure 10*). The effects of such disinformative data on hydrological inference were also discussed in Paper **IV**.

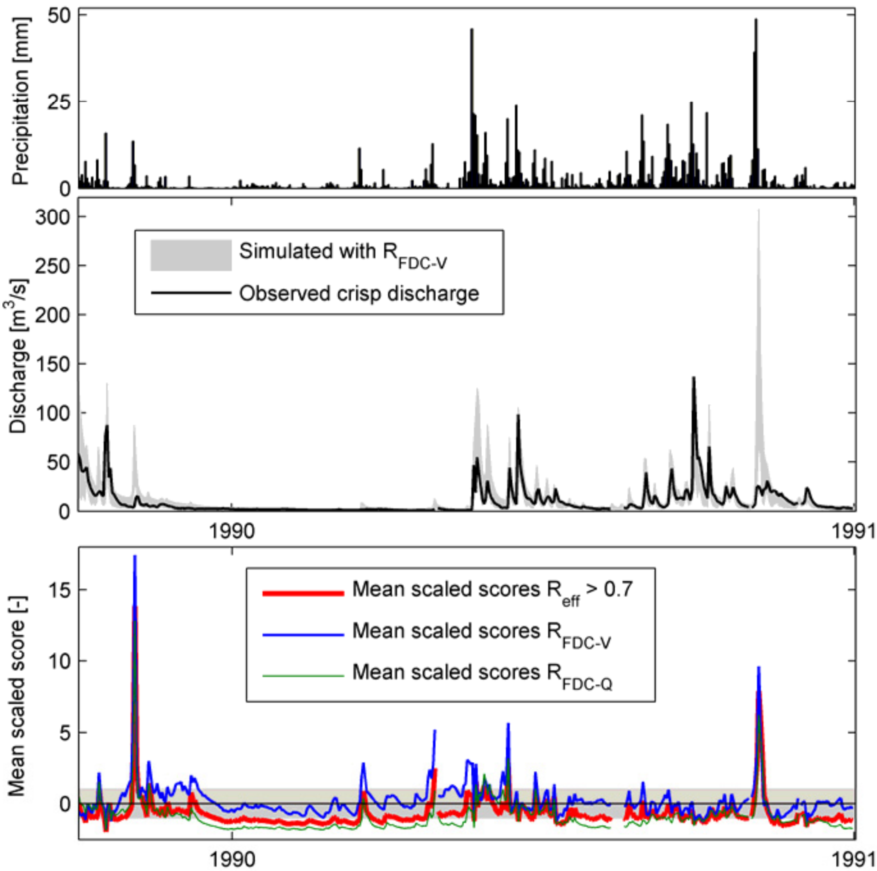

Figure 10. Daily precipitation in 1989–1990 (top) and predicted and observed crisp daily discharge for behavioural parameter-value sets from using R_{FDC-V} for calibration of WASMOD in the Paso La Ceiba catchment in 1980–1988 (middle). The mean scaled scores for all performance measures are plotted in the bottom plot where the grey area represents a scaled score from −1 to 1, i.e. a simulated discharge with a score inside this range is inside the discharge uncertainty limits.

Regionalisation of flow-duration curves

The R1 method resulted in a better overlap between the observed and predicted FDC than the R2 method as more of the uncertainty in the regionalisation was accounted for. This led to better overlap with observed ranges in the poorest cases but overestimation where the regionalisation worked the best (*Figure 11*). The use of 13 surrounding catchments provided a trade-off between overlap and overestimated uncertainty. The regionalisation worked least well for the four catchments with the most extreme FDC shapes in the dataset, e.g. Tamarindo that had the least base flow and a quick response to rainfall (*Figure 11*).

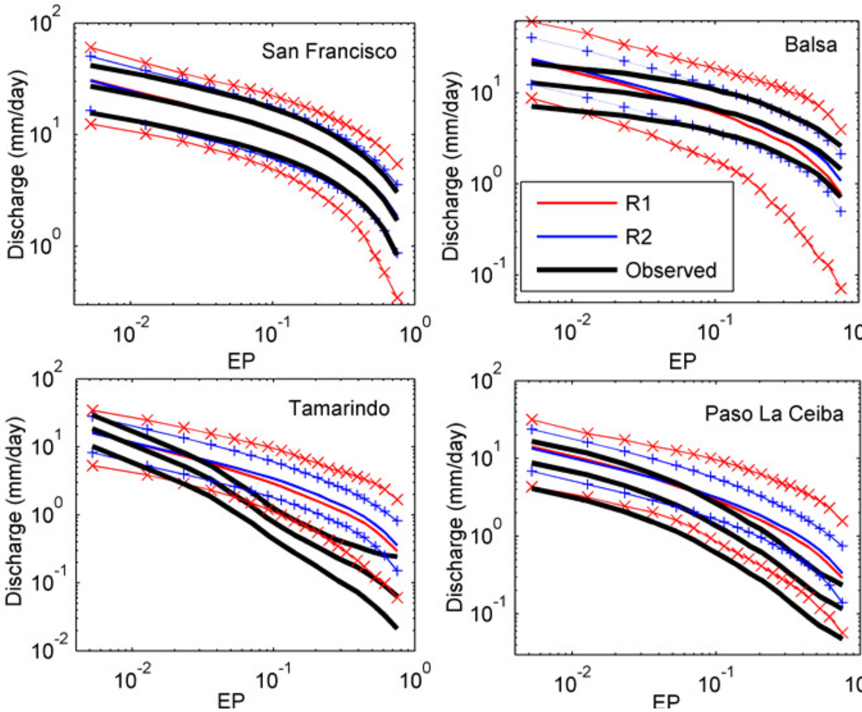

Figure 11. Examples of regionalised FDCs for the two regionalisation methods R1 and R2 compared to observed FDCs. Both discharge and EP exceedance percentage values are shown in log-space.

A regional water-balance model for Central America

When the regional model was calibrated using the locally calculated FDCs, simulations consistent with the observed FDCs were only found in 23 of the 32 basins. The basins with no behavioural simulations included four basins with differing run-off coefficients but similar mean annual precipitation (in northern Costa Rica), as well as the two Panamanian basins that deviated from the Budyko curve. This indicated that the CRN073 data were not sufficiently representative of the spatial precipitation variability in several parts of the region, and that this was an important reason for the lack of behavioural simulations in these basins. Compared to the simulations in the Paso La Ceiba catchment in Paper **III** where local precipitation data were used, the results with the regional precipitation data were less accurate (OP of 77% compared to 95%). Acceptable simulated results were only obtained in twelve basins in the regional simulations; there the observed and simulated discharge bounds overlapped for more than 50% of the time (for low, intermediate as well as high flows). A visual inspection showed that discharge peaks did not coincide with high amounts of precipitation in the preceding days in basins with lower high-flow OP values, compared to basins where this number was higher. To further test this hypothesis of unrepresentative precipitation data, the correlation between the current precipitation index (Smakhtin and Masse, 2000) and observed discharge for intermediate and high flows was compared to the high-flow OP values. It was seen that all basins with poor high-flow OP values also had low correlations between discharge and precipitation. The calibration to the regionalised FDCs was analysed in the 23 basins with behavioural results from the local FDC calibration. The R1-regionalisation generally resulted in more reliable simulations than the R2 method as there was a better overlap with the observed discharge, in particular for high flows (*Figure 12*). However, at a few stations, such as Paiwas (no. 18 in *Figure 12*), this was because of much greater predicted uncertainty, as compared to R2.

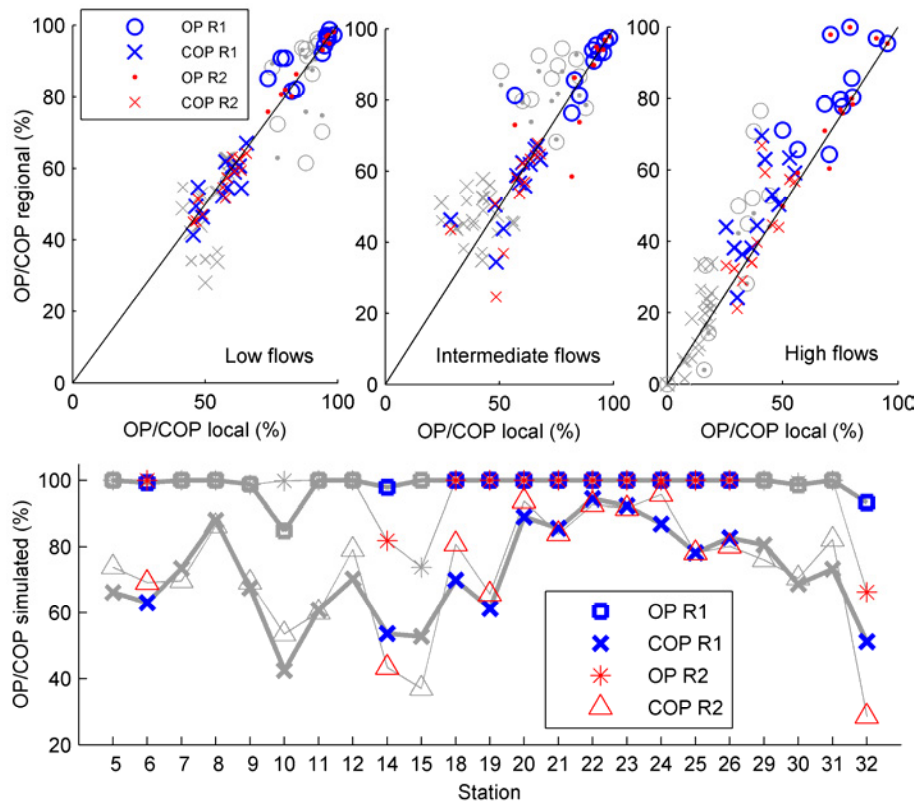

Figure 12. Upper panel: comparison of overlap between observed and simulated uncertainty bounds for simulations constrained with local and regionalised FDCs. Lower panel: comparison of regionalisation and local-calibration uncertainty bounds. OP and COP were calculated as the median of the overlap as a percentage of the locally-calibrated simulation bounds. The twelve basins where the local-calibration OP values were greater than 50% for low, intermediate and high flows are shown in red and blue.

The regionalisations generally resulted in reliable simulations compared to the local-data simulations. The greatest differences were seen in basins 14, 15 and 32 where the R2 method resulted in simulations that did not cover the low-flow range as well as for the R1 method, and in basin 10 where the R2-method resulted in poorer simulations for low flows (however, this basin had the poorest results from the local calibration). The reduction of the initial model uncertainty, i.e. the range between the maximum and minimum discharge bounds from all the 100,000 Monte Carlo runs, was in the mean 68% for R1 and 74% for R2. Where the regionalisation worked best it resulted in simulated discharge bounds that were just slightly wider than the locally-calibrated bounds (*Figure 13*).

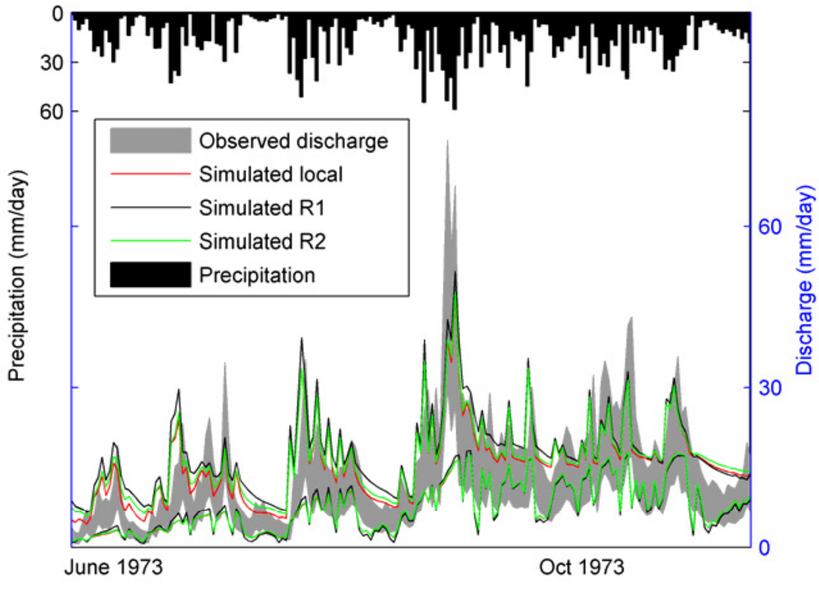

Figure 13. Comparison of simulated bounds from regionalisation and local calibration with observed discharge at San Francisco on the Santa Maria River (no. 24 in *Figure 12*).

Discussion

The methods that were developed in this thesis have shown potential for overcoming problems related to poor observational-data quality, including incomplete and fragmented time series and in the extreme case complete lack of discharge data. The last point, the prediction of discharge in ungauged basins, is an important prerequisite for a comprehensive mapping of water-resources that is needed for many types of management decisions. This work has not addressed numerical uncertainties in the mathematical implementation of model equations (Kavetski and Clark, 2010), and assessment of model-structural uncertainty through the use of multiple model structures.

Observational uncertainties were addressed in three ways in this thesis; quality control, quantitative estimation and development of model-evaluation techniques that addressed unquantifiable uncertainties. The easiest types of uncertainties to deal with are those data errors that can be detected and removed in a quality control. Such errors result from e.g. obvious measurement and digitisation errors. Detectable errors in precipitation, climate and discharge data were removed through quality control (Paper **I**–**III** and **V**). In Paper **I**, as much as 22% of the daily precipitation data had to be rejected because of poor quality. Quality problems have been detected even in high-quality datasets and have been shown to affect studies made with these (Viney and Bates, 2004). This is therefore not only a problem for developing regions like Central America.

Uncertainties that could be estimated from auxiliary data and/or expert opinion were then estimated. In this thesis such additional information was available only for discharge, were uncertainties were estimated through analyses of rating curves, temporal commensurability errors and realisation effects (Paper **II** and **V**). It is important to recognise that observational uncertainties may likely be affected by epistemic uncertainties and non-stationarity in addition to aleatory uncertainty, which will affect the information content of the data for hydrological inference (Paper **IV**). Such epistemic and non-stationary errors were e.g. found in the discharge-uncertainty estimations in Paper **II** and **V** which affected the choice of methods for dealing with these errors.

The most difficult uncertainties to handle are those epistemic errors that relate to lack of information such as unsampled variability and undetected quality problems. Such uncertainties cannot be quantitatively estimated (when additional information is lacking) but we may know something about the potential for such errors. For example, the potential for errors related to unsampled variability might be inferred from knowledge about the representativeness of the monitoring network in relation to the spatial and temporal variability of the hydro-meteorological variables. Some data uncertainties that were unknown prior to the modelling may also be revealed in a *post-hoc* analysis of the model residuals where it is possible to distinguish them from model-structural errors (Paper **III** and **V**). Because of the whole spectrum of possible errors in observational data, some data periods will be more informative than others in hydrologic inferences, while some periods may be completely disinformative. Two methods to deal with disinformative data were outlined in Paper **IV**. The first approach was to identify and remove events that were obviously disinformative prior to running the model, e.g. as evaluated on the basis of unrealistic runoff coefficients. The second approach was pursued in Paper **III**, i.e. to develop likelihood functions that are less sensitive to disinformative data. In this way epistemic errors may be accounted for implicitly.

Error assumptions and uncertainty estimation

How representative the estimated uncertainties are depends on the validity of the assumptions about the errors in the observed and simulated data. In Paper **I** the uncertainties in the interpolated precipitation data were found to be associated with substantial non-stationarity of an epistemic nature, especially in the spatial estimates. This was related to the temporally varying station density and the low number of gauges in relation to the high spatial and temporal variability of precipitation. The latter was especially true for the mountainous upper part of the basin that was used for modelling in Paper **III**. The regional-scale precipitation data in Paper **V** showed some obvious inhomogeneities that appeared to be related to the combined use of manual and automatic gauges, as well as the varying length of the gauge records (Magaña et al., 1999).

Discharge uncertainty was found to have non-stationary error characteristics, and epistemic types of uncertainties appeared important as a result of factors such as insufficient calibration of measurement equipment, poorly fitted rating curves, and un-sampled variability during periods that lacked ratings (Paper **II** and **V**). The non-stationary and complex characteristics of the uncertainties led to the choice of fuzzy and set-theoretic approaches for estimating these uncertainties in Paper **II** and for the modelling in Paper **III** and **V**. Fuzzy methods have previously been used in situations with few rating data for discharge-uncertainty estimation and model calibration (Krueger et al., 2010; Pappenberger et al., 2006). The discharge-uncertainty estimation presented in Paper **II** differed from these previous approaches; the predicted rating curves were not required to be contained within the data-uncertainty limits as the stage-discharge relationship was not stationary. The main assumptions in this method lied in the estimations of the uncertainties in the stage and discharge measurements. These were estimated as constant percentage errors, which is a simplification as measurements of discharge are more uncertain at low flows where cross-sectional irregularities have a larger effect on integration of flow area and relative uncertainties in current-meter measurements increase as velocity decrease (Pelletier, 1988; Petersen-Overleir et al., 2009). However, the estimation of ±25% uncertainty in discharge measurements appeared to be a reasonable estimate in this case as a similar figure was obtained for intermediate to high flows in the analysis of rating-curve residuals for the 35 Honduran stations in Paper **V**. The method was data-driven which means that the uncertainty estimates might not be representative for periods with unsampled variability, which is a limitation with this type of approach. Compared to a constant rating curve, there was substantial non-stationarity in the stage-discharge relationship with the largest relative differences occurring for low flows as a result of measurement uncertainties and non-stationarity of the river bed. The studies in Paper **II** and **V** confirmed the results of previous studies that have shown the sometimes large importance of non-stationary stage-discharge relationships (Jalbert et al., 2011; McMillan et al., 2010) and high uncertainty in discharge data (Di Baldassarre and Montanari, 2009; Petersen-Overleir et al., 2009). An important difference is the large number of ratings in the Honduran dataset in Paper **II** which permitted a data-driven approach.

The main assumption made for model prediction in Paper **III** and **V** was that the residual errors had a similar structure (in all its complexity) in calibration and prediction. This might not be a valid assumption for some types of predictions. If error characteristics are non-stationary the predictions should not be expected to bracket the observations. This has been the case for differential split-sample tests (Seibert, 2003). The non-stationarity of physical conditions over long time periods has been highlighted (Klemes, 1986a) – especially in recent years where human activities exert a large influence on the environment (Wagener et al., 2010). If information such as a FDC or recession curve is used for calibration, there might be an opportunity to include effects of estimated non-stationarities for future conditions in the uncertainty bounds for the information, e.g. similar to the realisation effect that was added to the FDC in Paper **V** to account for incomplete discharge-data records, or by using space-for-time proxies to estimate changed conditions (Wagener et al., 2010).

Model evaluation criteria and uncertainty estimation

One of the main advantages with the new method for model calibration (within the GLUE limits-of-acceptability framework) that was proposed in Paper **III** is that the simulated uncertainty bounds have a clear interpretation relative to the uncertainty in the observed FDC. Another advantage is the calibration to the whole range of flows and the possibility to use discharge data from another time period to overcome temporal mismatches of input and output data, while accounting for realisation uncertainty. A limitation on the sub-daily scale is that the timing of flow peaks was more uncertain than for traditional Nash-Sutcliffe calibration. Additional criteria will be required to constrain the timing where this uncertainty is not acceptable, as well as in catchments with snow and perhaps also for other types of models and hydrograph characteristics than those tested here.

Epistemic data uncertainties were addressed implicitly with this type of approach and this is an advantage of model calibration to information such as an FDC rather than error series, which are expected to be more sensitive to disinformative data. The extent to which the FDC-calibration is robust to disinformation needs to be addressed in further studies, but errors affecting the tails of the discharge distribution are clearly more important than those affecting the centre. Some previously unknown errors could be identified in the posterior analysis of the simulated discharge series and the scaled deviations to the limits of acceptability, and in some cases it was possible to disentangle probable model-structural errors from the effects of disinformative data events. In Paper **V** many basins had completely disinformative data as seen in the lack of behavioural simulations and low correlation between discharge and a current precipitation index. An FDC calibration without a posterior analysis is not recommended and the importance of this analysis was seen in Paper **V** as there were simulations consistent with the FDCs that had a poor overlap with the observed data. The main reason for this lack of fit was attributed to the non-representativeness of the regional precipitation data at the local scale in many catchments. This hypothesis was supported by the independent analysis of precipitation and discharge time series. Model-structural errors might be important in some cases if the parsimonious water-balance model used here is too simple. Testing of different model structures in future applications would provide more information in this regard.

Discharge uncertainty was quantified in these studies, but in many modelling applications there is no information at all about rating curves and data quality. The potential for such uncertainties must be kept in mind, as they can affect the results of hydrological analyses (McMillan et al., 2010).

Regionalisation using flow-duration curves

The FDC-calibration method was found to have potential for calibration to regionalised FDCs for ungauged basins and reduced the initial model uncertainty with around 70%. There were only minor differences to the local calibration in the basins where the prediction worked best. The R1 method resulted in better overlap with the observed data than R2 in the basins where the regionalisation of the FDCs worked less well. One advantage of using regionalised FDCs seems to be that there is an implicit water-balance constraint on the simulations, for example uncertainty in predicted flow peaks appeared to be smaller than those in the study by Yadav et al. (2007) where several different constraints were regionalised. It is however difficult to compare the results of studies made with different models and hydrological conditions. In some basins the regionalised bounds were much wider than the locally simulated bounds, and in these cases regionalisation of other constraints might further constrain the predicted uncertainty. Another strategy where this method is used for predictions in ungauged basins might simply be to make a few discharge measurements to constrain the predicted uncertainty further (Juston et al., 2009; Seibert and Beven, 2009). Regionalisations of FDCs have shown good results in nested (upstream/downstream) basins (Yu and Yang, 2000). This is perhaps not surprising given the dependency of the nested catchments, but it indicates that this FDC-calibration method could work well in such cases.

Observational uncertainties and water management

The methods that were developed in this thesis have provided means to overcome several problems related to scarce, poor-quality and non-existing data. Improved availability and quality of observational data is required to constrain predictive uncertainties and address epistemic uncertainties related to unmeasured variability. Sadly the opposite trend is seen in later years in many regions with reduced monitoring networks, for example in Honduras. Technical developments of new types of measurement equipment, such as low-cost wireless precipitation sensors and reliable direct discharge measurement devices may help to improve data availability and quality. However, technical solutions are not sufficient on their own, the value of observational data need to be recognised by government agencies to ensure adequate data management. This includes data quality control, data storage, security, and standardisation. The lack of representative precipitation data was identified as one of the most important limitations for regional water-balance modelling in Central America. The improvement of existing monitoring networks, such as in the Choluteca River basin in Honduras where 20% of the daily data had too poor quality, is also important for better use of existing resources. Well-maintained reliable automatic gauges with a dense spatial distribution that register precipitation at sub-hourly time scales are needed in this basin. Quality and availability are also important limitations when it comes to discharge data at the regional scale in Central America, where discharge data from several countries and authorities are unavailable for research.

In the context of uncertainty estimation and decision-making based on model predictions, it is important to also remember the potential for completely unknown unknowns that we know nothing about but which may affect the future.

Conclusions

This thesis presents methods for analysing, estimating and accounting for uncertainties in observational data in hydrological analysis and water-resources modelling. Most of the methods were tested in Central America, which is a region with a high variability of water resources in space and time and where data sometimes have poor quality. But many of these findings and problems are representative for other regions worldwide. Uncertainties in observations and modelled results were sometimes large, which manifest the importance of uncertainty analysis when it comes to deciding about water management. The methods for uncertainty analysis that were developed in this thesis provide a basis for such decisions and show both the potential and limitations of models for overcoming spatial and temporal gaps in data availability. Observational uncertainties were addressed in three ways; quality control for detectable errors, quantitative estimation where possible and development of model-evaluation techniques that addressed unquantifiable errors.

The time-variable rating-curve method allowed the estimation of the sometimes large temporal variability of stage-discharge relationships. Discharge uncertainty was found to have non-stationary error characteristics, and epistemic types of uncertainties appeared important as a result of factors such as insufficient calibration of measurement equipment, poorly fitted rating curves, and un-sampled variability.

The new FDC-calibration method enabled discharge uncertainty to be accounted for and this was also the basis for the rejection criterion. Compared to traditional lumped model-performance criteria this method resulted in a simultaneous calibration to the whole flow range. Additional criteria are needed to further constrain peak-flow timing at sub-daily time scales where this is of high importance, as well as for catchments with snow. The *post-hoc* analysis of the uncertainty in the simulated time series made it possible to identify potential model-structural errors and periods of disinformative data.

The method also showed potential for overcoming spatial and temporal discharge-data scarcity. It was possible to calibrate a model also for time periods with short discharge time series or where discharge was available for a non-overlapping period. Added realisation uncertainties from short or non-overlapping records can be estimated and added to the FDC in such cases. Two methods for regionalisation of FDCs were tested and these worked well in all but the most extreme types of flow regimes. Calibration to the regionalised FDCs reduced the initial model uncertainty by around 70%.

Precipitation data that were not representative of the high spatial and temporal variability was identified as the main obstacle to regional water-resources modelling in Central America. Precipitation data quality was also a problem in Honduras with 22% of the daily data for the Choluteca River basin having poor quality. The value of high-quality observational data for water-resources management and research needs to be recognised in this region. Data management including quality control, data storage and security, standardisation and availability for research need to be addressed in order to increase the knowledge about water-resources variability and improve water-resources management. This work has identified possible improvements in data management. It is also useful as a basis when analysing where and how measurements should be made to reduce uncertainties related to observational data.

Future research

This work has raised questions that are important to address. How robust is FDC calibration to different types of disinformative data? What are the advantages and disadvantages of the FDC for calibration compared to other types of information? How can snow parameters be constrained in this type of calibration? How can peak-flow timing at sub-daily time scales be further constrained? Testing of the FDC regionalisation in other places with precipitation data of higher quality, as well as other hydrological regimes, is also important for gaining further insights.

Acknowledgements

There are many people that have made these five years as a PhD-student to the great experience that it has been. First I would like to thank Sven, who as my main supervisor has been an invaluable support during these years and given me good advice in both the scientific and practical issues of my work. Special thanks for your excellent advice when it comes to writing. I've learned a lot from your structured approach. Jan, I'm very glad that you agreed to join my supervisor group, our discussions on methods and solutions for various problems over the years have taught me a lot about research and have been very important for my work. Thanks also for a good time visiting Zürich and a great day hiking to the top of a Swiss mountain! Thanks to Lars-Christer for good support and encouragement, especially in the first years of my PhD, and to Chong-Yu for your prompt and helpful advice with the many aspects of modelling and statistics. I also thank Sven, Lars-Christer and Chong-Yu for good times travelling in various parts of Central America. Thanks to Deliang for your support with my first paper and to Faruk who was my supervisor for only a short time, but was always enthusiastic and supportive of my work.

Keith, I first came in contact with your work (about topographic indices) when I did my master thesis. Little did I then think that I would get the opportunity to work with and have so many interesting discussions with you! Thanks for opening my eyes to the fascinating and intriguing world of red herrings and unknown unknowns, for your thoughtful advice and your encouragement.

José-Luis, you are a great friend and colleague. Thanks for all the support and cooperation in this work, for good discussions and for making the days in our shared office more fun. Thanks also for the warm hospitality that you and your family have shown me when visiting Honduras. Thanks to Lebing for your help with my research, for sharing ideas and implementations of those, for good discussions, and for good company on various conference trips over the years. Thanks to Michael for good cooperation and good times, as well as the opportunity to get some insights into the hydrology of the Victoria basin. Thanks to Alvaro for helping me with data for my regional modelling work, to Philip for helping me with the modelling on my third paper and to Alex for the good cooperation with my first paper. Thanks to Diana and Zairis for helping me with Honduran data, and to Anna for good discussions and for making both office and after-office hours more fun.

Special thanks to Allan whose inspiring teaching during my introductory course in hydrology sparked my interest for research in this field. It's been a joy to teach that same course to new students together with you over these years, and I've learned a lot about both teaching and hydrology from you in the process.

John, thanks for great discussions that have really helped my understanding of the topics we've been studying and that have made the research much more fun. Thanks also for good company and good times at various summer schools and a great trip from Scotland to England last year! Muchísimas gracias a Bea for the weeks I spent with you and your family in Costa Rica. I had a great time and really enjoyed the chance to be a "tica" for a while. And our sailing trip on the Pacific is a memory for life!

Thanks to present and former fellow hydrologists at the Earth Sciences Centre; Martin, Fritjof, Elin, Thomas, Agnes, José, Tito, Eduardo, Hanna, Roger, Kevin, Zhibing, Maria, Kristina, to Fredrik for a good advice when I started out as a PhD, and to all the meteorologists, glaciologists and other colleagues in Uppsala. Thanks to Tomas for excellent help and advice with computers as well as good discussions on books, and to Taher and Anna for excellent help in the printing office. Thanks to Dmitri, Thorsten and Jim for interesting discussions at various conferences.

The financial support from the Swedish Development Cooperation Agency, grant number 75007349, is gratefully acknowledged, together with the support for travels to international conferences by Liljewalchs resestipendium and Rektors resebidrag från Wallenbergsstiftelsen. I'm grateful to the staff at CIGEFI, University of Costa Rica for their kind assistance during my visit.

A great thing with having had two jobs during these years is that I've had the opportunity to work with many great colleagues also at IVL. Thanks to Sam for your support and encouragement, especially in starting the PhD. To Tony for good times and many interesting discussions about hydrology, I've learnt a lot from you and look forward to working together in the future. To the former and present colleagues in my department Micke, Helene, Erik, Marcus, Annika, Maria, Courtney, Faruk, Peter, Jonas, Mark, Micke, Magnus, Jessica, Charlotte, and all my other colleagues who make both the working days and Friday pubs interesting and fun!

Thanks to Jenny for many great times off from work, and Julia, Anna, Therese, Johan, Lotta, Jonas, Miriam, ... Thanks to all my kayaking friends for great times doing some very practical hydrological investigations of river and sea waves! Lastly I would like to thank all my family for their love and support, Mom and Dad, Henrik, Joel, Annette, Matilda, my beloved nieces and nephew Ellen, Benedikte and Erland, my Aunt and Uncle Kerstin and Björn and their family. Finally my Grandmothers Stina and Gunvor, who sadly cannot be there for my dissertation, but to whom this work is dedicated.

Sammanfattning på svenska (Summary in Swedish)

Observationsosäkerheter i vattenresursmodellering i Centralamerika - metoder för osäkerhetsuppskattning och modellutvärdering

De hydrologiska processernas variabilitet i både tid och rum har en stor påverkan på våra dagliga liv. Extrema händelser som översvämningar och torka kan leda till dödsfall samt förluster av egendom och försörjningsmöjligheter. Men även på en mindre skala påverkar vattenresursernas variabilitet många grenar av samhället såsom jordbruk, dricksvattenförsörjning, turism, infrastruktur och vattenkraft. Kunskap om hydrologiska processer samt den rumsliga och tidsmässiga variationen av de vattenresurser de skapar är därför basen för all typ av vattenförvaltning. Sådan kunskap är även av största vikt för att förebygga katastrofer och undvika konflikter i gränsöverskridande avrinningsområden. Grunden för sådan kunskap är observationsdata från mätningar av hydrologiska variabler. Sådana data är behäftade med osäkerheter som i vissa fall kan vara stora och om de inte beaktas kan de leda till felaktiga slutsatser från hydrologiska studier. Osäkerheter i data beror både på mätosäkerheter och på felaktig hantering. Tillgången till observationsdata är ofta otillräcklig i tid och rum och hydrologiska modeller är därför nödvändiga för en heltäckande analys av vattenresursers variabilitet.

Hydrologiska processer uppvisar en stor komplexitet i naturen; många sammanflätade energi-, vatten- och vegetationsprocesser pågår samtidigt på olika tids- och rumsskalor. Ett avrinningsområde är ett öppet system som kännetecknas av i hög grad ickelinjära processer och dåligt kända randvillkor. En sådan komplexitet är omöjlig att beskriva i en hydrologisk modell, som med nödvändighet blir en mycket förenklad beskrivning av verkligheten. Den förenklade processbeskrivningen, skillnader mellan modellens, processernas och observationsdatas upplösning, dåligt kända randvillkor, osäkerheter i observationsdata och numeriska beräkningsosäkerheter gör att en optimal, unik uppsättning av modellparametrar inte går att hitta vid modellkalibrering. Dessa osäkerheter behöver därför beaktas vid modellkalibrering och för att uppnå tillförlitliga resultat är en osäkerhetsanalys nödvändig. Osäkerhetsanalyser vid modellsimuleringar grundar sig på antaganden om hur de simulerade värdena avviker från de uppmätta och kan genomföras inom olika ramverk; Bayesisk statistik, oskarp logik och setteoretiska metoder, m.fl. Vilken typ av ramverk som lämpar sig bäst för en specifik studie beror på (den förväntade) komplexiteten hos modellfelen. När modellfelens struktur kan beskrivas med en statistisk modell lämpar sig Bayesiska statistiska metoder. När felens struktur uppvisar en större komplexitet och ickestationäritet lämpar sig ickeprobabilistiska eller possibilistiska metoder såsom setteoretiska metoder och sådana grundade på oskarp logik. De antaganden som gjorts om de olika ingående osäkerheterna och modellfelens struktur bör framgå tydligt och även testas i en utvärdering mot oberoende data. Osäkerheter i observationsdata har traditionellt sett inte beaktats vid modellkalibrering men har visat sig kunna påverka resultaten i hög grad. Två olika metoder för att avgöra modellresultatens samstämmighet med observationsdata kan urskiljas ur litteraturen; antingen baseras modellutvärderingskriteriet direkt på skillnaden mellan simulerade och uppmätta data eller på information, såsom en recessionskuva, beräknad från dessa. Den senare typen av modellkalibrering kan vara fördelaktig vid regionalisering av modellparametrar till avrinningsområden utan vattenföringsdata, eftersom informationen kan regionaliseras direkt och ofta har en starkare koppling till fysikaliska egenskaper hos avrinningsområdet än modellparametrarna.

De metoder som utvecklats i detta avhandlingsarbete har huvudsakligen tillämpats i Centralamerika men de problem med observationsosäkerheter som behandlats är även representativa för många andra områden i världen. Observationsosäkerheter hanterades på tre olika sätt i denna avhandling; kvalitetskontroll, kvantitativ uppskattning och utveckling av modellutvärderingsmetoder för beaktande av icke kvantifierbara osäkerheter. Det första steget var kvalitetskontroll av data för att avlägsna uppenbara felaktigheter. Vid kvalitetskontrollen av nederbördsdata i Cholutecaflodens avrinningsområde utarbetades metoder för att identifiera fyra typer av kvalitetsproblem som upptäcktes i data. Det resulterade i att hela 22 % av de dagliga data befanns ha dålig kvalitet. Nederbördsregimen uppvisade stor variabilitet i tid och rum, och stationsnätet befanns otillräckligt för en fullödig karaktärisering av tids- och rumsvariationerna.

Osäkerheter uppskattades sedan kvantitativt utifrån en dataanalys där tillräcklig information var tillgänglig. Osäkerheter i vattenföringsdata uppskattades dels vid beräkning av vattenföring med en oskarp regression för en tidsvariabel avbördningskurva, dels från en analys av redan beräknade avbördningskurvor med tillhörande data från 35 stationer i Honduras. Den relativa osäkerheten befanns i båda fallen vara störst vid låga flöden som ett resultat av högre mätosäkerheter samt större naturlig variabilitet, speciellt när vattendragets bottenprofil var ickestationär. Vid medelhöga och höga flöden var osäkerheten omkring ±25 % vid avbördningskurveanalysen för de 35 Honduranska stationerna. Detta överrensstämde med den uppskattning av osäkerhet i mätningen av vattenföring som användes för beräkningen med den tidsvariabla avbördningskurvan. I den senare analysen befanns tidsvariabiliteten vara stor jämfört med en konstant avbördningskurva, med skillnader på ±20 % för medelhöga och höga flöden samt ännu större skillnader för låga flöden (-60 till +90 %).

En metod för att hantera observationsosäkerheter vid modellkalibrering utvecklades inom den generella likelihood-osäkerhetsuppskattningsmetoden (GLUE). Metoden baserades på varaktighetskurvor som visar relationen mellan magnitud och frekvens av vattenföring. Acceptansgränser för den varaktighetskurva som beräknats från observerade data ansattes utifrån den uppskattning av osäkerheten i vattenföringsdata som tidigare gjorts. Modellutvärderingskriteriet innebar att alla simuleringar som resulterade i varaktighetskurvor som befann sig innanför acceptansgränserna för ett antal utvalda utvärderingspunkter på kurvan definierades som acceptabla. Valet av dessa utvärderingspunkter kan göras på olika sätt beroende på studiens mål och hydrografens karaktäristik. Två metoder testades i detta arbete. Den ena innebar att punkterna valdes utifrån jämna vattenföringsintervall på varaktighetskurvan. I den andra metoden valdes punkterna så att intervallen mellan punkterna representerade lika stor area under kurvan, vilket innebär att de representerar lika stor avrunnen volym vatten. Metoden testades för två olika modeller som kördes med olika tidsupplösning i två avrinningsområden. Jämfört med traditionell modellkalibrering med Nash-Sutcliffekriteriet befanns denna kalibreringsmetod ha ett antal fördelar; simuleringarna begränsades simultant för låga, medelhöga och höga flöden, acceptanskriteriet var inte godtyckligt utan grundades på osäkerheten i utvärderingsdata, modellkalibrering var möjlig när indata och jämförelsedata inte överlappar i tiden och kriteriet verkade vara mindre känsligt för disinformativa data. Ett antal fall där kriteriet inte är tillräckligt på egen hand identifierades också; beräkningar där hög säkerhet i tidpunkten för flödestoppen är viktigt vid modellering med hög tidsupplösning (här timupplösning), samt vid beräkningar i avrinningsområden med snö. En *post hoc* analys användes för att analysera överrensstämmelsen mellan de observerade och simulerade tidsserierna. En sådan analys är ett viktigt tredje steg vid hantering av observationsosäkerheter i hydrologisk modellering eftersom den gör det möjligt att upptäcka disinformativa dataperioder samt troliga modellstrukturfel där dessa två typer av fel går att särskilja från varandra. En sådan analys är även av högsta vikt för att granska antagandena om modellfelens struktur samt dessas stationäritet mellan olika tidsperioder.

Varaktighetskurvor kan regionaliseras, dvs. predikteras för områden där vattenföringsdata saknas. Två metoder baserade på viktade linjära kombinationer av data från de hydrologiskt sett mest liknande avrinningsområdena testades och befanns fungera väl, förutom för de områden som hade de mest extrema flödesfördelningarna. En regional modell för Centralamerika sattes upp och kalibrerades med hjälp av de regionaliserade varaktighetskurvorna. I de områden där detta fungerade bäst var skillnaden mot kalibrering med lokala data mycket liten. Prediktionerna var fortfarande tillförlitliga men mycket osäkrare, för de områden där varaktighetskurveregionaliseringen fungerade sämre. I genomsnitt minskades den initiala modellosäkerheten med 70 % med hjälp av denna regionaliseringsmetod.

Det största hindret för en uttömmande regional analys av vattenresursernas förekomst i tid och rum i Centralamerika befanns vara bristen på tillförlitliga nederbördsdata. Endast i en tredjedel av de studerade områdena var överrensstämmelsen mellan nederbörd och avrinning acceptabel. Vattenföringsdatas brist på tillgänglighet och stundom låga kvalitet utgör också ett problem, men detta kan i många fall hanteras med de metoder som utvecklats i detta avhandlingsarbete. De metoder och analyser som ingår i denna avhandling gör det möjligt att uppskatta osäkerheter som finns i hydrologiska data och modeller. De kan därmed utgöra viktiga redskap vid vattenresursförvaltning.

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., Oconnell, P.E. and Rasmussen, J., 1986. An Introduction to the European Hydrological System - Systeme Hydrologique Europeen, She .1. History and Philosophy of a Physically-Based, Distributed Modeling System. Journal of Hydrology, 87(1-2): 45-59.

Aguilar, E., Peterson, T.C., Obando, P.R., Frutos, R., Retana, J.A., Solera, M., Soley, J., Garcia, I.G., Araujo, R.M., Santos, A.R., Valle, V.E., Brunet, M., Aguilar, L., Alvarez, L., Bautista, M., Castanon, C., Herrera, L., Ruano, E., Sinay, J.J., Sanchez, E., Oviedo, G.I.H., Obed, F., Salgado, J.E., Vazquez, J.L., Baca, M., Gutierrez, M., Centella, C., Espinosa, J., Martinez, D., Olmedo, B., Espinoza, C.E.O., Nunez, R., Haylock, M., Benavides, H. and Mayorga, R., 2005. Changes in precipitation and temperature extremes in Central America and northern South America, 1961-2003. Journal of Geophysical Research-Atmospheres, 110(D23): D23107, doi: 10.1029/2005JD006119.

Alexandersson, H., 2003. Korrektion av nederbörd enligt enkel klimatologisk metodik. (Correction of precipitation with a simple climatological approach). No. 111, SMHI Reports Meteorology. Swedish Meteorological and Hydrological Institute, Norrköping.

Alfaro, E.J., 2002. Some characteristics of the precipitation annual cycle in Central America and their relationships with its surrounding tropical oceans. Tópicos Meteorológicos y Oceanográficos, 9(2): 88-103.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration – guidelines for computing crop water requirements, FAO Irrigation and Drainge Paper 56. FAO, Rome.

Amador, J.A., 2003. Climate and climate variability in the Arenal river basin of Costa Rica. In: Climate and Water: Transboundary Challenges in the Americas. vol. Kluwer Academic Publishers Dordrecht, pp. 317-349.

Amador, J.A., Alfaro, E.J., Lizano, O.G. and Magana, V.O., 2006. Atmospheric forcing of the eastern tropical Pacific: A review. Progress in Oceanography, 69(2-4): 101-142.

Aronica, G.T., Candela, A., Viola, F. and Cannarozzo, M., 2006. Influence of rating curve uncertainty on daily rainfall-runoff model predictions. In: Predictions in Ungauged Basins: Promise and Progress. vol., pp. 116–124.

Balairón Pérez, L., Álvarez Rodríguez, J., Borrell Brito, E. and Delgado Sánchez Sánchez, M., 2004. Balance hídrico de Honduras - documento principal. CEDEX, Madrid.

Beven, K.J., 1989. Changing Ideas in Hydrology - the Case of Physically-Based Models. Journal of Hydrology, 105(1-2): 157-172.

Beven, K.J., 1993. Prophecy, Reality and Uncertainty in Distributed Hydrological Modeling. Advances in Water Resources, 16(1): 41-51.

Beven, K.J., 2001. Rainfall-Runoff Modelling - The Primer. Wiley, Chichester.

Beven, K.J., 2002. Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. Hydrological Processes, 16(2): 189-206.

Beven, K.J., 2006. A manifesto for the equifinality thesis. Journal of Hydrology, 320(1-2): 18-36.

Beven, K.J., 2009. Environmental Modelling: An Uncertain Future? Routledge, London.

Beven, K.J., 2010. Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models Preface. Hydrological Processes, 24(12): 1537-1547.

Beven, K.J. and Binley, A., 1992. The Future of Distributed Models - Model Calibration and Uncertainty Prediction. Hydrological Processes, 6(3): 279-298.

Beven, K.J. and Freer, J., 2001. A dynamic TOPMODEL. Hydrological Processes, 15(10): 1993-2011.

Beven, K.J., Smith, P.J. and Freer, J.E., 2008. So just why would a modeller choose to be incoherent? Journal of Hydrology, 354(1-4): 15-32.

Blazkova, S. and Beven, K., 2009. A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. Water Resources Research, 45: W00B16, doi: 10.1029/2007WR00672.

Bonacci, O., 1999. Water circulation in karst and determination of catchment areas: example of the River Zrmanja. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 44(3): 373-386.

Box, G.E.P., 1976. Science and Statistics. Journal of the American Statistical Association, 71(356): 791-799.

Box, G.E.P., 1979. Robustness in the strategy of scientific model building. In: Robustness in Statistics. vol. Academic Press New York.

Boyle, D.P., Gupta, H.V. and Sorooshian, S., 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. Water Resources Research, 36(12): 3663-3674.

Brath, A., Montanari, A. and Toth, E., 2004. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. Journal of Hydrology, 291(3-4): 232-253.

Budyko, M.I., 1974. Climate and life. Academic press, London.
Bulygina, N., McIntyre, N. and Wheater, H., 2009. Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis. Hydrology and Earth System Sciences, 13(6): 893-904.
Bulygina, N., McIntyre, N. and Wheater, H., 2011. Bayesian conditioning of a rainfall-runoff model for predicting flows in ungauged catchments and under land use changes. Water Resources Research, 47: W02503, doi: 10.1029/2010wr009240.
Burkham, D.E. and Dawdy, D.R., 1970. Error analysis of streamflow data for a alluvial stream, Geological Survey professional paper, 655-C. USGS, Washington, D.C.
Burn, D.H., 1990a. An Appraisal of the Region of Influence Approach to Flood Frequency-Analysis. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 35(2): 149-165.
Burn, D.H., 1990b. Evaluation of Regional Flood Frequency-Analysis with a Region of Influence Approach. Water Resources Research, 26(10): 2257-2265.
Buytaert, W. and Beven, K., 2009. Regionalization as a learning process. Water Resources Research, 45: W11419, doi: 10.1029/2008wr007359.
Carter, R.W., 1970. Accuracy of current meter measurements. Paper presented at International Symposium on Hydrometry, Koblenz. International Association of Hydrological Sciences: Publication No. 99, 86-98 pp.
Choi, H.T. and Beven, K.J., 2007. Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in an application of TOPMODEL within the GLUE framework. Journal of Hydrology, 332(3-4): 316-336.
Clark, M.P. and Slater, A.G., 2006. Probabilistic quantitative precipitation estimation in complex terrain. Journal of Hydrometeorology, 7(1): 3-22.
Clarke, R.T., 1973. A review of some mathematical models used in hydrology, with observations on their calibration and use. Journal of Hydrology, 19: 1-20.
Clarke, R.T., 1999. Uncertainty in the estimation of mean annual flood due to rating-curve indefinition. Journal of Hydrology, 222(1-4): 185-190.
Crawford, N.H. and Linsley, R.K., 1966. Digital simulation in hydrology: Stanford Watershed Model IV. 39, Techical Report 39. Department of Civil Engineering, Stanford University CA.
Criss, R.E. and Winston, W.E., 2008. Do Nash values have value? Discussion and alternate proposals. Hydrological Processes, 22(14): 2723-2725.
Di Baldassarre, G. and Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrology and Earth System Sciences, 13(6): 913-921.

Diaz, H.F., Hoerling, M.P. and Eischeid, J.K., 2001. ENSO variability, teleconnections and climate change. International Journal of Climatology, 21(15): 1845-1862.

Dubois, D. and Prade, H., 1980. Fuzzy Sets and Systems: Theory and Applications. Academic Press, San Diego, CA.

Eischeid, J.K., Baker, C.B., Karl, T.R. and Diaz, H.F., 1995. The Quality-Control of Long-Term Climatological Data Using Objective Data-Analysis. Journal of Applied Meteorology, 34(12): 2787-2795.

Enfield, D.B. and Alfaro, E.J., 1999. The dependence of Caribbean rainfall on the interaction of the tropical Atlantic and Pacific oceans. Journal of Climate, 12(7): 2093-2103.

Engeland, K., Xu, C.Y. and Gottschalk, L., 2005. Assessing uncertainties in a conceptual water balance model using Bayesian methodology. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 50(1): 45-63.

Fekete, B.M., Vorosmarty, C.J. and Grabs, W., 2002. High-resolution fields of global runoff combining observed river discharge and simulated water balances. Global Biogeochemical Cycles, 16(3): 1042, doi: 10.1029/1999gb001254.

Feng, S., Hu, Q. and Qian, W.H., 2004. Quality control of daily meteorological data in China, 1951-2000: A new dataset. International Journal of Climatology, 24(7): 853-870.

Ferson, S., Joslyn, C.A., Helton, J.C., Oberkampf, W.L. and Sentz, K., 2004. Summary from the epistemic uncertainty workshop: consensus amid diversity. Reliability Engineering & System Safety, 85(1-3): 355-369.

Flambard, O., 2003. Rapport Actividad 1 - Inventario, análisis y diagnostico de la red hydrometeorologica, ref D4168/RAP/2003-00059-B (limited access). EDF, Toulouse.

Freeze, R.A. and Harlan, R.L., 1969. Blueprint for a physically-based, digitally-simulated hydrologic response model. Journal of Hydrology, 9: 237–258.

Genereux, D.P., Wood, S.J. and Pringle, C.M., 2002. Chemical tracing of interbasin groundwater transfer in the lowland rainforest of Costa Rica. Journal of Hydrology, 258(1-4): 163-178.

George, R.K., Waylen, P. and Laporte, S., 1998. Interannual variability of annual streamflow and the Southern Oscillation in Costa Rica. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 43(3): 409-424.

Gong, L., Halldin, S. and Xu, C.Y., 2011. Global-scale river routing - an efficient time-delay algorithm based on HydroSHEDS high-resolution hydrography. Hydrological Processes, 25: 1114-1128. doi: 10.1002/hyp.7795.

Gonzalez-Rouco, J.F., Jimenez, J.L., Quesada, V. and Valero, F., 2001. Quality control and homogeneity of precipitation data in the southwest of Europe. Journal of Climate, 14(5): 964-978.

Gottschalk, L. and Askew, A., 1987. Hydrology and data acquisition. In: Hydrology 2000. vol. 171. IAHS Publ., pp. 79-90.

Grayson, R.B., Moore, I.D. and Mcmahon, T.A., 1992. Physically Based Hydrologic Modeling .2. Is the Concept Realistic. Water Resources Research, 28(10): 2659-2666.

GRDC, 2010. Global Runoff Data Centre. Global Runoff Data Centre. http://grdc.bafg.de. Accessed: 2010-02-23. Koblenz, Germany

Gupta, H.V., Beven, K.J. and Wagener, T., 2005. Model calibration and uncertainty estimation. In: Encyclopedia of Hydrological Sciences. vol. John Wiley & Sons, Ltd. Chichester.

Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1-2): 80-91.

Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4): 751-763.

Gupta, H.V., Wagener, T. and Liu, Y.Q., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrological Processes, 22(18): 3802-3813.

Helton, J.C. and Oberkampf, W.L., 2004. Alternative representations of epistemic uncertainty. Reliability Engineering & System Safety, 85(1-3): 1-10.

Herschy, R.W., 1970. The Evaluation of Errors at Flow-Measurement Stations. Paper presented at International Symposium on Hydrometry, Koblenz. International Association of Hydrological Sciences: Publication no. 99, 109-131 pp.

Hojati, M., Bector, C.R. and Smimou, K., 2005. A simple method for computation of fuzzy linear regression. European Journal of Operational Research, 166(1): 172-184.

Holländer, H.M., Blume, T., Bormann, H., Buytaert, W., Chirico, G.B., Exbrayat, J.F., Gustafsson, D., Hölzel, H., Kraft, P., Stamm, C., Stoll, S., Blöschl, G. and Flühler, H., 2009. Comparative predictions of discharge from an artificial catchment (Chicken Creek) using sparse data. Hydrology and Earth System Sciences, 13(11): 2069-2094.

Holmes, M.G.R., Young, A.R., Gustard, A. and Grew, R., 2002. A region of influence approach to predicting flow duration curves within ungauged catchments. Hydrology and Earth System Sciences, 6(4): 721-731.

Hornberger, G.M. and Spear, R.C., 1981. An Approach to the Preliminary-Analysis of Environmental Systems. Journal of Environmental Management, 12(1): 7-18.

Houghton-Carr, H.A., 1999. Assessment criteria for simple conceptual daily rainfall-runoff models. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 44(2): 237-261.

Huard, D. and Mailhot, A., 2008. Calibration of hydrological model GR2M using Bayesian uncertainty analysis. Water Resources Research, 44(2): W02424, doi: 10.1029/2007wr005949.
Isaaks, E.H. and Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York, 561 pp.
Jakeman, A.J. and Hornberger, G.M., 1993. How Much Complexity Is Warranted in a Rainfall-Runoff Model. Water Resources Research, 29(8): 2637-2649.
Jalbert, J., Mathevet, T. and Favre, A.C., 2011. Temporal uncertainty estimation of discharges from rating curves using a variographic analysis. Journal of Hydrology, 397(1-2): 83-92.
Juston, J., Seibert, J. and Johansson, P.O., 2009. Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. Hydrological Processes, 23(21): 3093-3109.
Kavetski, D. and Clark, M.P., 2010. Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. Water Resources Research, 46: W10511, doi: 10.1029/2009wr008896.
Kavetski, D., Fenicia, F. and Clark, M., 2011. Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment. Water Resources Research. doi: 10.1029/2010WR009525, in press.
Keesman, K. and Vanstraten, G., 1990. Set Membership Approach to Identification and Prediction of Lake Eutrophication. Water Resources Research, 26(11): 2643-2652.
Kirchner, J.W., 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research, 42(3): W03S04, doi: 10.1029/2005wr004362.
Klemes, V., 1986a. Dilettantism in Hydrology - Transition or Destiny. Water Resources Research, 22(9): 177S-188S.
Klemes, V., 1986b. Operational Testing of Hydrological Simulation-Models. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 31(1): 13-24.
Krause, P., Boyle, D.P. and Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, 5: 89-97.
Krueger, T., Freer, J., Quinton, J.N., Macleod, C.J.A., Bilotta, G.S., Brazier, R.E., Butler, P. and Haygarth, P.M., 2010. Ensemble evaluation of hydrological model hypotheses. Water Resources Research, 46: W07516, doi: 10.1029/2009WR00784.
Kuczera, G., 1996. Correlated rating curve error in flood frequency inference. Water Resources Research, 32(7): 2119-2127.

Lehner, B., Verdin, K. and Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos, Transactions, AGU, 89(10): 93-94.

Liu, Y., Freer, J., Beven, K.J. and Matgen, P., 2009. Towards a limits of acceptability approach to the calibration of hydrological models: extending observation error. Journal of Hydrology, 367: 93-103.

Lopes, V.L., 1996. On the effect of uncertainty in spatial distribution of rainfall on catchment modelling. Catena, 28(1-2): 107-119.

Magaña, V., Amador, J.A. and Medina, S., 1999. The midsummer drought over Mexico and Central America. Journal of Climate, 12(6): 1577-1588.

Maurer, E.P., Adam, J.C. and Wood, A.W., 2009. Climate model based consensus on the hydrologic impacts of climate change to the Rio Lempa basin of Central America. Hydrology and Earth System Sciences, 13(2): 183-194.

McIntyre, N., Lee, H., Wheater, H., Young, A. and Wagener, T., 2005. Ensemble predictions of runoff in ungauged catchments. Water Resources Research, 41(12): W12434, doi: 10.1029/2005wr004289.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T. and Clark, M., 2010. Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. Hydrological Processes, 24(10): 1270-1284.

McMillan, H., Jackson, B., Clark, M., Kavetski, D. and Woods, R., 2011. Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. Journal of Hydrology, 400(1-2): 83-94.

Michaud, J.D. and Sorooshian, S., 1994. Effect of Rainfall-Sampling Errors on Simulations of Desert Flash Floods. Water Resources Research, 30(10): 2765-2775.

Montanari, A., 2007. What do we mean by 'uncertainty"? The need for a consistent wording about uncertainty assessment in hydrology. Hydrological Processes, 21(6): 841-845.

Montanari, A. and Toth, E., 2007. Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins? Water Resources Research, 43(5): W05434, doi: 10.1029/2006wr005184.

Moyeed, R.A. and Clarke, R.T., 2005. The use of Bayesian methods for fitting rating curves, with case studies. Advances in Water Resources, 28(8): 807-818.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models 1. A discussion of principles. Journal of Hydrology, 10: 282-290.

Obled, C., Wendling, J. and Beven, K., 1994. The Sensitivity of Hydrological Models to Spatial Rainfall Patterns - an Evaluation Using Observed Data. Journal of Hydrology, 159(1-4): 305-333.

Oreskes, N., Shraderfrechette, K. and Belitz, K., 1994. Verification, Validation, and Confirmation of Numerical-Models in the Earth-Sciences. Science, 263(5147): 641-646.

Pappenberger, F., Matgen, P., Beven, K.J., Henry, J.B., Pfister, L. and Fraipont de, P., 2006. Influence of uncertain boundary conditions and model structure on flood inundation predictions. Advances in Water Resources, 29(10): 1430-1449.

Parajka, J., Merz, R. and Bloschl, G., 2005. A comparison of regionalisation methods for catchment model parameters. Hydrology and Earth System Sciences, 9(3): 157-171.

Paturel, J.E., Servat, E. and Vassiliadis, A., 1995. Sensitivity of Conceptual Rainfall-Runoff Algorithms to Errors in Input Data - Case of the Gr2m Model. Journal of Hydrology, 168(1-4): 111-125.

Pelletier, P., 1988. Uncertainties in the single determination of river discharge: a literature review. Canadian Journal of Civil Engineering, 15(5): 834-850.

Peña, M. and Douglas, M.W., 2002. Characteristics of wet and dry spells over the Pacific side of Central America during the rainy season. Monthly Weather Review, 130(12): 3054-3073.

Petersen-Overleir, A., 2004. Accounting for heteroscedasticity in rating curve estimates. Journal of Hydrology, 292(1-4): 173-181.

Petersen-Overleir, A., 2006. Modelling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 51(3): 365-388.

Petersen-Overleir, A., Soot, A. and Reitan, T., 2009. Bayesian Rating Curve Inference as a Streamflow Data Quality Assessment Tool. Water Resources Management, 23(9): 1835-1842.

Portig, W.H., 1976. The climate of Central America. In: World Survey of Climatology. vol. 12, Climates of Central and South America. Elsevier New York, pp. 405-464.

Refsgaard, J.C. and Knudsen, J., 1996. Operational validation and intercomparison of different types of hydrological models. Water Resources Research, 32(7): 2189-2202.

Refsgaard, J.C., van der Sluijs, J.P., Brown, J. and van der Keur, P., 2006. A framework for dealing with uncertainty due to model structure error. Advances in Water Resources, 29(11): 1586-1597.

Reichl, J.P.C., Western, A.W., McIntyre, N.R. and Chiew, F.H.S., 2009. Optimization of a similarity measure for estimating ungauged streamflow. Water Resources Research, 45: W10423, doi: 10.1029/2008wr007248.

Reitan, T. and Petersen-Overleir, A., 2008. Bayesian power-law regression with a location parameter, with applications for construction of discharge rating curves. Stochastic Environmental Research and Risk Assessment, 22(3): 351-365.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M. and Franks, S.W., 2010. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resources Research, 46: W05521, doi: 10.1029/2009wr008328.

Ross, J.L., Ozbek, M.M. and Pinder, G.F., 2009. Aleatoric and epistemic uncertainty in groundwater flow and transport simulation. Water Resources Research, 45: W00b15, doi: 10.1029/2007wr006799.

Rougier, J.C., 2010. Formal Bayes Methods for Model Calibration with Uncertainty. In: Applied Uncertainty Analysis for Flood Risk Management. vol. Imperial College Press / World Scientific London.

Schaefli, B. and Gupta, H.V., 2007. Do Nash values have value? Hydrological Processes, 21(15): 2075-2080.

Schmidt, A.R., 2002. Analysis of stage-discharge relations for open-channel flow and their associated uncertainties. Ph. D. Thesis, University of Illinois, Urbana.

Seibert, J., 1999. Regionalisation of parameters for a conceptual rainfall-runoff model. Agricultural and Forest Meteorology, 98-9: 279-293.

Seibert, J., 2001. On the need for benchmarks in hydrological modelling. Hydrological Processes, 15: 1063-1064.

Seibert, J., 2003. Reliability of model predictions outside calibration conditions. Nordic Hydrology, 34(5): 477-492.

Seibert, J. and Beven, K.J., 2009. Gauging the ungauged basin: how many discharge measurements are needed? Hydrology and Earth System Sciences, 13(6): 883-892.

Sevruk, B., 1986. Correction of precipitation measurements, summary report. In: Correction of precipitation measurements. vol. 23. ETH, Geographisches Institut Zurich, pp. 13-23.

Shrestha, R.R., Bardossy, A. and Nestmann, F., 2007. Analysis and propagation of uncertainties due to the stage-discharge relationship: a fuzzy set approach. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 52(4): 595-610.

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S. and Zehe, E., 2003. IAHS decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 48(6): 857-880.

Skop, E. and Loaiciga, H.A., 1998. Investigating catchment hydrology and low flow characteristics using GIS. Nordic Hydrology, 29(2): 105-128.

Smakhtin, V.Y. and Masse, B., 2000. Continuous daily hydrograph simulation using duration curves of a precipitation index. Hydrological Processes, 14(6): 1083-1100.

Son, K. and Sivapalan, M., 2007. Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. Water Resources Research, 43(1): W01415, doi: 10.1029/2006wr005032.

Spear, R.C. and Hornberger, G.M., 1980. Eutrophication in Peel Inlet .2. Identification of Critical Uncertainties Via Generalized Sensitivity Analysis. Water Research, 14(1): 43-49.
Stedinger, J.R., Vogel, R.M., Lee, S.U. and Batchelder, R., 2008. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resources Research, 44: W00b06, doi: 10.1029/2008wr006822.
Stephens, G.L. and Kummerow, C.D., 2007. The remote sensing of clouds and precipitation from space: A review. Journal of the Atmospheric Sciences, 64(11): 3742-3765.
Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S.W. and Srikanthan, S., 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resources Research, 45: W00b14, doi: 10.1029/2008wr006825.
Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S. and Gupta, H.V., 2003. Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. Hydrological Processes, 17(2): 455-476.
Wagener, T. and Montanari, A., 2011. Convergence of Approaches towards Reducing Uncertainty in Predictions in Ungauged Basins (PUB). Water Resources Research: doi: 10.1029/2010WR009469, in press.
Wagener, T., Sivapalan, M., Troch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Kumar, P., Rao, P.S.C., Basu, N.B. and Wilson, J.S., 2010. The future of hydrology: An evolving science for a changing world. Water Resources Research, 46: W05301, doi: 10.1029/2009wr008906.
Vandewiele, G.L. and Elias, A., 1995. Monthly Water-Balance of Ungauged Catchments Obtained by Geographical Regionalization. Journal of Hydrology, 170(1-4): 277-291.
Waylen, P. and Laporte, M.S., 1999. Flooding and the El Nino-Southern Oscillation phenomenon along the Pacific coast of Costa Rica. Hydrological Processes, 13(16): 2623-2638.
Weedon, G., Gomes, S., Viterbo, P., Österle, H., Adam, J., Bellouin, N., Boucher, O. and Best, M., 2010. The WATCH forcing data 1958-2001: a meteorological forcing dataset for land surface- and hydrological-models, WATCH Technical Report No. 22.
Whitehead, P. and Young, P., 1979. Water-Quality in River Systems - Monte-Carlo Analysis. Water Resources Research, 15(2): 451-459.
Villarini, G. and Krajewski, W.F., 2010. Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall. Surveys in Geophysics, 31(1): 107-129.
Wilson, C.B., Valdes, J.B. and Rodriguez-Iturbe, I., 1979. On the Influence of the Spatial-Distribution of Rainfall on Storm Runoff. Water Resources Research, 15(2): 321-328.
Viney, N.R. and Bates, B.C., 2004. It never rains on Sunday: The prevalence and implications of untagged multi-day rainfall accumulations in the

Australian high quality data set. International Journal of Climatology, 24(9): 1171-1192.

Winsemius, H.C., Schaefli, B., Montanari, A. and Savenije, H.H.G., 2009. On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. Water Resources Research, 45: W12422, doi: 10.1029/2009wr007706.

Vogel, R.M. and Fennessey, N.M., 1994. Flow-Duration Curves. 1: New Interpretation and Confidence-Intervals. Journal of Water Resources Planning and Management-Asce, 120(4): 485-504.

Xu, C.-Y., 2002. WASMOD - The water and snow balance modeling system. In: Mathematical Models of Small Watershed Hydrology and Applications. vol. Water Resources Publications LLC Highlands Ranch, Colorado, U.S, pp. 555-590.

Xu, C.Y., 2003. Testing the transferability of regression equations derived from small sub-catchments to a large area in central Sweden. Hydrology and Earth System Sciences, 7(3): 317-324.

Yadav, M., Wagener, T. and Gupta, H., 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. Advances in Water Resources, 30(8): 1756-1774.

Yang, C., Chandler, R.E., Isham, V.S. and Wheater, H.S., 2006. Quality control for daily observational rainfall series in the UK. Water and Environment Journal, 20(3): 185-193.

Yang, J., Reichert, P., Abbaspour, K.C. and Yang, H., 2007. Hydrological modelling of the chaohe basin in china: Statistical model formulation and Bayesian inference. Journal of Hydrology, 340(3-4): 167-182.

Yilmaz, K.K., Gupta, H.V. and Wagener, T., 2008. A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. Water Resources Research, 44(9): W09417, doi: 10.1029/2007wr006716.

You, J.S., Hubbard, K.G., Nadarajah, S. and Kunkel, K.E., 2007. Performance of quality assurance procedures on daily precipitation. Journal of Atmospheric and Oceanic Technology, 24(5): 821-834.

Younger, P.M., Freer, J.E. and Beven, K.J., 2009. Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework. Hydrological Processes, 23(14): 1988-2003.

Yu, P.S. and Yang, T.C., 2000. Using synthetic flow duration curves for rainfall-runoff model calibration at ungauged sites. Hydrological Processes, 14(1): 117-133.

Zadeh, L.A., 2005. Toward a generalized theory of uncertainty (GTU) - an outline. Information Sciences, 172(1-2): 1-40.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 833

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-152074

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2011